# DOCUMENT RESUME

ED 480 146 EA 032 713

| | |
|---|---|
| AUTHOR | Hite, Seven J. |
| TITLE | Reviewing Quantitative Research To Inform Educational Policy Processes. Fundamentals of Educational Planning. |
| INSTITUTION | United Nations Educational, Scientific, and Cultural Organization, Paris (France). International Inst. for Educational Planning. |
| REPORT NO | No-69 |
| ISBN | ISBN-92-803-1210-3 |
| PUB DATE | 2001-00-00 |
| NOTE | 120p. |
| AVAILABLE FROM | United Nations Educational, Scientific, and Cultural Organization, International Inst. for Educational Planning, 7-9 rue Eugene-Delacroix, 75116 Paris, France. Web site: http://www.unesco.org/ iiep. |
| PUB TYPE | Books (010) -- Guides - Non-Classroom (055) -- Reports - Evaluative (142) |
| EDRS PRICE | EDRS Price MF01 Plus Postage. PC Not Available from EDRS |
| DESCRIPTORS | Credibility; Educational Assessment; *Educational Planning; Educational Policy; *Educational Research; Elementary Secondary Education; Policy Analysis; Policy Formation; Research Design; Research Skills; *Research Tools; *Research Utilization |
| IDENTIFIERS | Research Implementation; Research Synthesis |

ABSTRACT

Educational planners and policymakers are rarely able to base their decision-making on sound information and research, according to this book. Because the situation is even more difficult in developing countries, educational policy often is based on research conducted in others parts of the world. This book provides a practical framework that can be used to locate and evaluate quantitative research. It supplies information on the types of sources to be consulted, the research tools available, and the processes involved in analyzing research findings. The first chapter addresses issues in reviewing educational research and the process of reviewing research. The second chapter focuses on locating research documents and provides information on print and online sources. Chapter 3 describes reliability and validity in analyzing quantitative research. Chapter 4 discusses components of research documents for purposes of analysis. Chapter 5 looks at ways of evaluating research and determining research methodology type. The final chapter discusses synthesizing and presenting research analysis results, including citation and referencing according to the APA format, graphic presentation of results, and drawing conclusions across studies. Appended are a glossary of common Internet terms and a short list of educational Web sites. (WFA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY HAS BEEN GRANTED BY

J. Hall

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Fundamentals of Educational Planning

ED 480 146

# Reviewing quantitative research to inform educational policy processes

Steven J. Hite

UNESCO:
International
Institute
for Educational
Planning

EA 032 713

BEST COPY AVAILABLE

Included in the series:*

2. The relation of educational plans to economic and social planning, *R. Poignant*
4. Planning and the educational administrator, *C.E. Beeby*
5. The social context of educational planning, *C.A. Anderson*
6. The costing of educational plans, *J. Vaizey, J.D. Chesswas*
7. The problems of rural education, *V.L. Griffiths*
8. Educational planning; the adviser's role, *A. Curle*
9. Demographic aspects of educational planning, *Ta Ngoc C.*
10. The analysis of educational costs and expenditure, *J. Hallak*
11. The professional identity of the educational planner, *A. Curle*
12. The conditions for success in educational planning, *G.C. Ruscoe*
13. Cost-benefit analysis in educational planning, *M. Woodhall*
18. Planning educational assistance for the second development decade, *H.M. Philips*
20. Realistic educational planning, *K.R. McKinnon*
21. Planning education in relation to rural development, *G.M. Coverdale*
22. Alternatives and decisions in educational planning, *J.D. Montgomery*
23. Planning the school curriculum, *A. Lewy*
24. Cost factors in planning educational technological systems, *D.T. Jamison*
25. The planner and lifelong education, *P. Furter*
26. Education and employment: a critical appraisal, *M. Carnoy*
27. Planning teacher demand and supply, *P. Williams*
28. Planning early childhood care and education in developing countries, *A. Heron*
29. Communication media in education for low-income countries, *E.G. McAnany, J.K. Mayo*
30. The planning of nonformal education, *D.R. Evans*
31. Education, training and the traditional sector, *J. Hallak, F. Caillods*
32. Higher education and employment: the IIEP experience in five less-developed countries, *G. Psacharopoulos, B.C. Sanyal*
33. Educational planning as a social process, *T. Malan*
34. Higher education and social stratification: an international comparative study, *T. Husén*
35. A conceptual framework for the development of lifelong education in the USSR, *A. Vladislavlev*
36. Education in austerity: options for planners, *K. Lewin*
37. Educational planning in Asia, *R. Roy-Singh*
38. Education projects: elaboration, financing and management, *A. Magnen*
39. Increasing teacher effectiveness, L.W. Anderson
40. National and school-based curriculum development, *A. Lewy*
41. Planning human resources: methods, experiences and practices, *O. Bertrand*
42. Redefining basic education for Latin America: lessons to be learned from the Colombian Escuela Nueva, *E. Schiefelbein*
43. The management of distance learning systems, *G. Rumble*
44. Educational strategies for small island states, *D. Atchoarena*
45. Judging educational research based on experiments and surveys, *R.M. Wolf*
46. Law and educational planning, *I. Birch*
47. Utilizing education and human resource sector analyses, *F. Kemmerer*
48. Cost analysis of educational inclusion of marginalized populations, *M.C. Tsang*.
49. An efficiency-based management information system, *W. W. McMahon*.
50. National examinations: design, procedures and reporting, *J. P. Keeves*.
51. Education policy-planning process: an applied framework, *W. D. Haddad, with the assistance of T. Demsky*
52. Searching for relevance: the development of work orientation in basic education, *Wim Hoppers*
53. Planning for innovation in education, *Dan E. Inbar*
54. Functional analysis (management audits) of the organization of ministries of education, *R. Sack and M. Saïdi*
55. Reducing repetition: issues and strategies, *Thomas Owen Eisemon*
56. Increasing girls and women's participation in basic education, *Nelly P. Stromquist*
57. Physical facilities for education: what planners need to know, *John Beynon*
58. Planning learner-centred adult literacy programmes, *Susan E. Malone and Robert F. Arnove*
59. Training teachers to work in schools considered difficult, *Jean-Louis Auduc*
60. Evaluating higher education, *Jeanne Lamoure Rontopoulou*
61. The shadow education system: private tutoring and its implication for planners, *Mark Bray*
62. School-based management, *Ibtisam Abu-Duhou*
63. Globalization and educational reform: what planners need to know, *Martin Carnoy*
64. Decentralization of education: why, when, what and how?, *N. McGinn and T. Welsh*
65. Early childhood education: need and opportunity, *D. Weikart*
66. Planning for education in the context of HIV/AIDS, *Michael J. Kelly*
67. Legal aspects of educational planning and administration, *Claude Durand-Prinborgne*
68. Improving school effectiveness, *Jaap Scheerens*

* Also published in French. Other titles to appear.

BEST COPY AVAILABLE

# Reviewing quantitative research to inform educational policy processes

Steven J. Hite

Paris 2001
UNESCO: International Institute for Educational Planning

The Swedish International Development Co-operation Agency (Sida)
has provided financial assistance for the publication of this booklet.

Published in 2001 by the United Nations
Educational, Scientific and Cultural Organization
7 place de Fontenoy, F 75352 Paris 07 SP
Printed in France by SAGIM
Cover design by Pierre Finot

ISBN 92-803-1210-3
© UNESCO 2001

# Fundamentals of educational planning

The booklets in this series are written primarily for two types of clientele: those engaged in educational planning and administration, in developing as well as developed countries; and others, less specialized, such as senior government officials and policy-makers who seek a more general understanding of educational planning and of how it is related to overall national development. They are intended to be of use either for private study or in formal training programmes.

Since this series was launched in 1967 practices and concepts of educational planning have undergone substantial change. Many of the assumptions which underlay earlier attempts to rationalize the process of educational development have been criticized or abandoned. Even if rigid mandatory centralized planning has now clearly proven to be inappropriate, this does not mean that all forms of planning have been dispensed with. On the contrary, the need for collecting data, evaluating the efficiency of existing programmes, undertaking a wide range of studies, exploring the future and fostering broad debate on these bases to guide educational policy and decision-making has become even more acute than before. One cannot make sensible policy choices without assessing the present situation, specifying the goals to be reached, marshalling the means to attain them and monitoring what has been accomplished. Hence planning is also a way to organize learning: by mapping, targeting, acting and correcting.

The scope of educational planning has been broadened. In addition to the formal system of education, it is now applied to all other important educational efforts in non-formal settings. Attention to the growth and expansion of education systems is being complemented and sometimes even replaced by a growing concern for the quality of the entire educational process and for the control of its results. Finally, planners and administrators have become more and more aware of the importance of implementation strategies and of the role of different regulatory mechanisms in this respect: the choice of financing methods, the examination and certification procedures or various other regulation and incentive structures. The concern of planners is twofold: to reach a better understanding of the validity of education in its own empirically observed specific dimensions and to help in defining appropriate strategies for change.

The purpose of these booklets includes monitoring the evolution and change in educational policies and their effect upon educational planning requirements; highlighting current issues of educational planning and analyzing them in the context of their historical and societal setting; and disseminating methodologies of planning which can be applied in the context of both the developed and the developing countries.

For policy-making and planning, vicarious experience is a potent source of learning: the problems others face, the objectives they seek, the routes they try, the results they arrive at and the unintended results they produce are worth analysis.

In order to help the Institute identify the real up-to-date issues in educational planning and policy-making in different parts of the world, an Editorial Board has been appointed, composed of two general editors and associate editors from different regions, all professionals of high repute in their own field. At the first meeting of this new Editorial Board in January 1990, its members identified key topics to be covered in the coming issues under the following headings:

1. Education and development.
2. Equity considerations.
3. Quality of education.
4. Structure, administration and management of education.
5. Curriculum.
6. Cost and financing of education.
7. Planning techniques and approaches.
8. Information systems, monitoring and evaluation.

Each heading is covered by one or two associate editors.

The series has been carefully planned but no attempt has been made to avoid differences or even contradictions in the views expressed by the authors. The Institute itself does not wish to impose any official doctrine. Thus, while the views are the responsibility of the authors and may not always be shared by UNESCO or the IIEP, they warrant attention in the international forum of ideas. Indeed, one of the purposes of this series is to reflect a diversity of experience and opinions by giving different authors from a wide range of backgrounds and disciplines the opportunity of expressing their views on changing theories and practices in educational planning.

Educational planners and policy-makers are faced with a number of problems when defining and implementing policies: one of these is to locate reliable information to back their decision-making. A review of the research literature that already exists can be most useful in weighing up the different options available.

When undertaking a quantitative research review, it is important not only to be well acquainted with the basic principles of quantitative analysis, such as reliability and validity, but also to have mastered certain fundamental research techniques.

This booklet takes the reader through all the steps involved in reviewing quantitative research literature, from relatively simple tasks such as locating the most relevant documents, to the more delicate matter of evaluating and synthesizing the literature in hand. Educational policy-makers at all levels, as well as planners and researchers, will find it an invaluable aid.

Steven Hite, Associate Professor of Educational Research Theory and Methodology at Brigham Young University, has managed to present this complex subject in a clear and accessible way. The IIEP would like to thank him for his noteworthy contribution.

Gudmund Hernes
Director, IIEP

## *Composition of the Editorial Board*

*Chairman:*
Gudmund Hernes
Director, IIEP

*General Editors:*
Françoise Caillods
IIEP

T. Neville Postlethwaite (Professor Emeritus)
University of Hamburg
Germany

*Associate Editors:*
Jean-Claude Eicher
University of Bourgogne
France

Claudio de Moura Castro
Inter-American Development Bank
USA

Kenneth N. Ross
IIEP

Richard Sack
Executive Secretary
Association for the Development
of Education in Africa (ADEA)
France

Rosa Maria Torres
International Consultant
Argentina

# Preface

Quality education is increasingly recognized as playing a vital role in the progress of society, at all levels. In order to further development and growth, nations are dedicating more financial resources than ever before to goals such as the universalization of primary education, and the expansion and improvement of the secondary and tertiary systems. In this context, efficiency and effectiveness are key issues, and education systems must prove capable of standing up to scrutiny concerning the measurable personal and social benefits deriving from their new commitments.

At the heart of the issue regarding the actual outcomes of educational investment, rests the question of what reliable and valid evidence exists to support or assess the efficacy of these revitalized educational endeavours. While the potential usefulness of educational research in the policy process is not in question, how to proceed in acquiring and applying the best research data available remains a challenging, and often seemingly insurmountable, task. Major obstacles in this area are, on the one hand, the perception that high levels of technical, statistical and methodological expertise are required and, on the other, the lack of a reasonably direct and systematic approach for locating and evaluating available educational research.

In this booklet, Steven Hite presents a framework that can be used by educational policy-makers, governmental ministry personnel, educational research specialists and indeed any involved and interested educator, for the identification and evaluation of educational research results. Even those with little technical training and expertise can begin to access and make a critical appraisal of the available educational research literature base on topics of importance, whether these concern the local, regional, national, or global level.

In practical fashion the author guides his reader through the process of locating relevant research reports, including how to make optimal use of valuable tools, such as computers, the Internet and inter-library loan systems. He then details the type of systematic analysis that must be undertaken in order to establish the quality and usefulness of the findings, and gives a thorough description of the different methods and methodologies that may be encountered in the literature. One of the points he raises is that for various reasons it is often not possible to locate high-quality educational research, and that it is therefore important to be able to make the best possible use of even incomplete or flawed research reports.

It is clear that a careful and methodical review of the available research literature can potentially yield significant evidence regarding the utility and outcomes of educational policy and programmes. The dependable and replicable techniques presented in this booklet for the location, evaluation and synthesis of educational research should be useful to any interested and responsible educational specialist.

Kenneth Ross
Associate Editor

# Contents

Preface 11

I. Introduction 15
  Policy purposes for reviewing research 15
  Quantitative focus of research review approach 16
    Issues in reviewing educational research 17
  Process of research review 18

II. Locating research documents 19
  Key words and terms 20
  Local sources 24
    Public libraries 24
    School and university libraries 26
    Government archives 27
  Research handbooks and encyclopedias 28
  Inter-library loan 33
    Online Computer Library Centre (OCLC) 33
    'Paper' requests for ILL 34
  Online and Internet sources 34
    The World Wide Web 35
    Combining search terms 35
    Subscription services 38
    Free services 39
    University libraries online 42

III. Analyzing quantitative research 45
  Reliability and validity 45
    Reliability 45
    Validity 49

IV. Components of research documents for purposes of analysis 53
  Purpose and problem 55
  Research methods and methodologies 58

General procedures 60
Sampling 62
Measurement 65
Data analysis 68
Results 71
Conclusions and generalizability 72

V. Evaluating research 75
Determining research methodology type 75
Statistical procedures approach 76
Design characteristics approach 79
Descriptive research 82
Correlational research 86
Causal research 90
'Gaps' in educational research 95

VI. Synthesizing and presenting research analysis results 97
Synthesis versus serial recitation 97
Narrative presentation 98
Essential v. important v. interesting information 101
Building an argument: the inverted triangle 102
Citation and referencing – APA format 105
Graphic presentation 107
Tables 107
Figures 108
Drawing conclusions across studies 109
Presenting both sides - positive and negative 110
Accounting for limitations 110
Balancing technical limitations with policy potential 111

Conclusion 113

References 115

Appendices 119

# I. Introduction

The assertion that good educational policy is more likely to be developed if the process is informed by quality educational research is generally accepted. However, research and policy literature generally affirm that much educational research is not of particularly high quality. Thus systematic, dependable and replicable ways of locating, evaluating and synthesizing high-quality educational research to inform the development of educational policy are desperately needed. Where high-quality research is not available, the limitations of lower quality research need to be identified, examined and contained in the use of research in the policy process. The following chapters suggest a way forward in this important area.

## *Policy purposes for reviewing research*

Formation and implementation of policy are inherently political endeavours. The policy process in education, therefore, is often influenced negatively by changes in political conditions and trends (Randall, Cooper and Hite, 1999). For example, the election of a new government with a 'new' educational agenda may restrict or discontinue a line of research and investigation before it has sufficient time to reach fruition. If governments change often, and priorities for educational research change with them, sustaining a sophisticated and productive research agenda will be almost impossible. Even less desirable, the basis for a particular educational policy and associated educational research initiatives may be based solely on current political necessity or expediency (securing votes or proving short-term solutions).

In a context of political fluctuation in educational imperatives – which is commonplace in certain developing areas of the world – the most dependable research basis for formation and implementation of educational policy may need to be built upon a review of previous quantitative, scientifically sound, research conducted in other parts of the world.

Creating and building policy upon a review of educational research literature can accomplish any or all of the following four objectives:

1. The design of a particular educational policy or the proposal for investigation of educational policy should be informed by the best research available;
2. The results of current policies and policy initiatives can be evaluated through comparison with educational research conducted in other settings;
3. The methods and procedures used in researching the appropriate development of policy in other countries, settings or times can guide the creation of a research initiative in a reasonably comparable setting;
4. The analysis and/or evaluation of a current or past policy can be guided by work that has been done in other locations or at other times.

Thus, reviewing the research can provide a solid foundation on which to build, develop and sustain educational policy. To a significant degree, the form and structure of the research literature review will be influenced by whichever of these four purposes is central to the policy endeavour. However, the underlying issues and procedures for evaluating educational research reports are the same for all.

## *Quantitative focus of research review approach*

The field of educational research is much more complex today than even 20 years ago. Research theory and methodology from fields such as sociology and anthropology have generated non-traditional approaches and methodologies that are now being used in many educational settings. Most traditional and non-traditional approaches to educational research are grouped under the general headings of *quantitative* and *qualitative* research. Various authors have credibly described and explored the history and development of the fields of quantitative and qualitative research in education. In this book, however, no such attempts will be made. This work will assume that research being sought to inform decisions impacting the formation of educational policy is quantitative. This position is based on several assertions:

1. The majority of educational research of the last 50 years is primarily quantitative in design. Access to a broader set of research reports, in more diverse chronological and geographical circumstances, is therefore more likely in both traditional (for example, physical 'paper and book' libraries and archives) and contemporary (for example, computer-based) search methods.

2. The *lingua franca* of policy and of the policy formation process at virtually all levels (local, state/province/district, regional, national, international) is predominately quantitative. To speak with professional adequacy and effectiveness in this context requires the adoption and use of the prevailing expectations and language of quantitative reasoning.

3. Discourse and investigation in institutions that are involved globally in education research, policy and funding (for example, UNESCO, the World Bank, and regional banks such as the IDB and the ADB) are based in the quantitative paradigm. Interacting with these types of agencies and accessing their resources is contingent on facility in their mode of operation.

4. The objective of quantitative research, as distinguished from qualitative research, has always been to assert, investigate, and generate results and theories that are generalizable. Quantitative research in education was modelled on the scientific method basic to the physical sciences. Therefore, the results of quantitative educational research present the greatest likelihood of applicability in settings other than those in which they were conducted.

### *Issues in reviewing educational research*

In beginning the process of reviewing research literature, one might assume that it is necessary to be an expert in statistics and research design in order to produce a credible, defensible, and usable review. Such expertise is not a precondition, but the review must proceed methodically, with careful attention to details that might not appear important in a more casual reading of research material (Fink, 1998; Katzer, Cook and Crouch, 1998; Pyrczak, 1999).

The process of reviewing research focuses on eight basic elements:

1. Purpose and problem;
2. Research methodology;
3. General procedures;
4. Sampling;
5. Measurement;
6. Data analysis;
7. Results;
8. Conclusions and generalizability.

These issues will be woven into the fabric of this book in many of the different sections. Each issue will be presented in depth in Chapter 3: Analyzing quantitative research.

## *Process of research review*

There are basically four steps in creating a review of research literature:

1. Locating the research reports;
2. Analyzing each report's research design elements;
3. Evaluating the quality and usefulness of each report;
4. Presenting and synthesizing the results of that analysis and evaluation.

More detailed descriptions of how to proceed with each of these steps will be presented in the following chapters. Additionally, a short bibliography of reference materials will be provided as a resource for more extensive presentations of each of these steps.

# II. Locating research documents

Locating research documents (for example, government reports, professional journal articles, books reporting original research) of appropriate quality or usefulness represents the first major challenge in the process of using educational research to inform policy formation. While modern technology certainly offers the potential of broadening access to much more information than in the past, technology is not necessarily the only way, or even at times the best way, to acquire needed resources. A careful combination of traditional and contemporary approaches to locating and acquiring research documents is usually the most effective method for building a workable base of research data.

Before dealing with the different sources of research literature for review, a point must be made (and it will be re-stated many times hereafter for emphasis): *The LEAST successful reviewers will see libraries, archives, books and web-based documents and sources as the most productive sources of information; but the MOST successful reviewers will understand that these sources only point to the most important resources – PEOPLE.*

The use of the following sources and techniques will only be fully productive if the research reviewer understands the primary task of these tools as identifying the people and organizations that are conducting or have conducted research on the topic. Then the successful reviewer will contact these people and sources directly (via telephone, the Internet, or traditional mail) to develop an active network of researchers and organizations dealing with the subject.

The process of locating research documents begins with identifying key words and terms, followed by use of local sources, interlibrary loan sources, and online and Internet sources of research documents.

## *Key words and terms*

Prior to searching any source for research documents, the first step is to build a list of 'key' words and terms that will guide the search, whether a library card catalogue or a computer database is being used. While the list will likely change during the search process, time spent refining key words and terms before beginning will significantly increase efficiency.

The first step in determining key words and terms is to be able to clearly articulate the research question(s) in very precise form. *Figure 2.1* contains some examples of poor and better research questions, with the key words and terms they produce.

**Figure 2.1 Research topic: Reading curriculum**

**Poor research question:**

Which reading curriculum works best?

*Potential key words and terms:*

*Words:* reading

*Terms:* reading curriculum, best reading curriculum

**Better research question:**

Which early primary school reading curricula increase the level of comprehension, recall and abstract reasoning in both boys and girls?

*Potential key words and terms:*

*Words:* reading, comprehension, recall, reasoning, gender, boys, girls

*Terms:* reading curriculum, primary reading curriculum, primary school reading curriculum, level of comprehension, level of recall, level of abstract reasoning, abstract reasoning
The terms *gender*, *boys* and *girls* could be added to any of these terms for more precision and access to a broader base of literature.

**Note:** Question #2 is preferable because it indicates interest in more specific functions and outcomes of reading programmes. It also is preferable to Question #1 because it defines 'best' with specific performance criteria. Question #2 also takes into account a common 'confounding' variable in early reading curriculum – gender. Clearly the words and terms generated by Question #2 will more likely produce results when searching for usable research literature, if for no other reason than it provides more access points to the literature through a larger set of relevant words and terms.

While key words and terms appear clear to the person choosing them, a number of aspects needs to be considered in their use. The definition and use of words and terms are subject to time and location. For instance, the term *gender* might not generate all of the available information on the topic because the term *sex* has been and continues to be used in many settings to refer to the same subject. But the term *sex* can also refer, obviously, to a much wider set of applications than the more specific term 'gender'.

**Figure 2.2 Research topic – Student assessment**

**Poor research question:**

What is the best way to assess students?

*Potential key words and terms:*

*Words:* assess, student(s)

*Terms:* student assessment

**Better research question:**

What reliable and valid instruments are available for assessing student academic achievement?

*Potential key words and terms:*

*Words:* assess, assessment, reliability, validity, instrument(s), student(s), achievement

*Terms:* student assessment, reliable student assessment, valid student assessment, reliable instruments, valid instruments, reliable assessment instruments, valid assessment instruments, academic achievement, student academic achievement, student achievement, assessing academic achievement, assessing student academic achievement, assessing academic achievement.

**Note:** Again, Question #2 is preferable here because it indicates that both instruments and student performance are of particular importance and interest. New researchers may often be reluctant to become more specific in constructing their research questions because they fear that specificity will limit the range of their inquiry. This example illustrates that specificity actually creates more access points to the literature than Question #1.

Sensitivity to the effect of time on meaning and use of terms is also critical. If the example of reading curriculum had focused on the influence of race and ethnicity, rather than gender, and if the location of the search was the USA, interesting challenges would have occurred. The first column of *Table 2.1* shows some of the current racial and ethnic categories, as used in the USA Population Census 2000. The second column lists some of the terms that have been used in the research literature for the same categories, or more specific sub-categories, during the last 40 years. If the research literature review were intended to include literature that has been created over the last 40 years or that uses 'non-standard' terms and categories, each of these terms would have to be used to produce a comprehensive search.

**Table 2.1 Selected USA Census 2000, racial and ethnic categoriescompared to previously used historical terms**

| Selected USA Census 2000 categories | Sample of other historical or contemporary terms |
|---|---|
| Hispanic or Latino | Hispanic<br>Latino<br>Chicano<br>Cuban<br>Puerto Rican<br>South American<br>Spanish<br>Mexican<br>Mexican-American |
| Black or African-American | Black<br>African-American<br>Afro-American<br>Negro<br>African<br>American of African Descent |
| American Indian or Alaska Native | American Indian<br>Native American<br>AmerIndian<br>Indian<br>West Indian<br>Western Indian<br>Eskimo<br>Aleut |

Each of these terms, including variations for specific groups or sub-groups, needs to be carefully considered and appropriately used. In addition to the racial and ethnic terms shown in this example, many other types of words and terms may be subject to the same type and range of variation. Such variations may reflect political movements and shifts, legal decisions, merging of groups or clarification of group differences, or a myriad of other possibilities. Whatever the reason, limited access to a significant segment of research literature may result if the reviewer does not critically consider the potential influence of historical, political or social conditions on the use and meaning of key words and terms.

## *Local sources*

While electronic databases are increasing both in number and in ease of access, much can be gained by first exhausting local sources of research. Local sources are the most likely to have research directly relevant to the specific setting. Thus the reviewer is more likely to begin with material that is usable and valid for the context; the search can then be extended to include a wider range of research literature that will be carefully analyzed for local applicability and generalizability. Traditionally, relevant local sources include public libraries, school and university libraries, and government archives. However, the strongest local sources of information are not the traditional 'archival' locations. Rather, the most potent local resource is actually PEOPLE, particularly those who are actively involved in the educational and policy issues at hand. While a number of different traditional local sources will be considered below, a prudent researcher will always maximize and leverage the human knowledge resources available in the local setting above all other research sources.

### *Public libraries*

Public libraries typically house research information that may be somewhat general in nature but can be useful in research reviews (Mann, 1987). Common sources of useful library information would be newspaper archives, international magazines (such as Time, The Economist, Newsweek) and specialty sections focusing on government documents and professional literature.

While these sources are not considered to be the best resources for high-quality research, they often provide references to larger, more professional research work, including the names of people and organizations conducting that research. Therefore, items located in public libraries should be considered primarily as indicators of who should be contacted directly for more extensive, higher quality work.

A second function of local library materials is use for historical reference. Since library materials may have been on the shelves for a number of years, they typically do not represent either the newest information or researchers currently conducting studies, although they can still provide important information including access to individuals and organizations still functioning in the research arena. As historical documents, however, library holdings are valuable as they provide important information on past educational research initiatives, as well as a cumulative record of results and trends in education and educational research.

In local libraries, the central means for locating information is the card catalogue. Most people attempting a review of research literature will be familiar with the use of card catalogues. However, following a few basic guidelines will increase the likelihood that relevant items can be found if they are a part of the library's collection.

Typically three cards in the catalogue represent library items: one alphabetized by the last name of the primary author, one alphabetized by the primary subject of the item, and sometimes one alphabetized by title.

Knowing that any one item can be located by author, subject or title should suggest the use of a process of linking or bridging between the card being viewed and other potential listings based on information on the card, such as:

1. If an article is found by using the primary subject listing, then the card catalogue should also be consulted for other listings by the same author.

2. If an article is found by using the author listing, also check in the card catalogue under the primary subjects indicated in the key words and terms section of the card. These key words and terms on the catalogue cards can help to expand your own initial list of key words and terms derived from your question(s).

Using one source to lead to another, i.e. author search to indicate additional subjects or subject search to identify new authors, is one of the most efficient ways of 'building' a search. Always consider any one source as a way of leverage into other materials, rather than seeing each article or report as an end in itself.

The card catalogue provides other 'clues' for extending your search. Just as researchers tend to continue work in a specific area of inquiry, or expand into related areas, publishers and journals often focus their efforts in certain areas of interest. For example, if a publisher has printed one book on the use of student assessment instruments in Kenyan education, then this publisher is highly likely to have published other books and materials appropriate to this same theme. Publishers and journals are more than happy to provide publication and author lists when requested. After all, providing such lists enables them to sell more materials, which is what they are in business to do! So never hesitate to contact publishers through mail, telephone, or e-mail, that have shown to be in the business of publishing material in your specific area of interest.

### *School and university libraries*

School and university libraries typically provide the best research literature of the traditional library sources. The quickest way to determine availability and quality of local educational research documents is to visit the library of the best accessible university. If the resources of that university library are limited in terms of volume, quality and/or recency, then it is unlikely that research literature in any other locally available source (library or otherwise) will yield much of use.

The great advantage of university libraries over public libraries is that the function of the university is to train students and to make the best research in their field available to them. This does not guarantee that good and current research will be in the library, only that the library materials will be the best available. Also, university libraries typically have 'discipline specialists' who are trained to direct patrons easily and quickly to the available research literature. These specialists would also likely know professors, research units etc. at the university that could further help you in your area of inquiry.

Conducting research in a university library is similar to researching in public libraries. A common difference, however, is that university libraries generally have a more direct connection to government resources and to research documents generated at the local, national and international levels. Therefore, as is the case with public libraries, the single most valuable resource represented by university libraries (particularly in developing contexts) is in locating government units conducting or responsible for generating relevant research.

### *Government archives*

Most government units are required to generate reports of their activities and, often, evaluations of the programmes falling under their authority. In fact, a researcher who can gain access to government documents has found perhaps the best available information regarding domestic research. The fact that these documents often represent the best available information, however, does not ensure that the information is necessarily reliable or valid. The same evaluative criteria must be applied to these documents as are used with other research sources.

In many countries, government archives are readily accessible to the public, with the records arranged and indexed in a manner similar to that of regular libraries. Frequently, in fact, government documents are available at the larger and more prestigious university libraries, making access to government archive information quite simple. In other countries, however, permission to access government records must typically be sought from the director of the unit that actually conducts the research for a particular subject. Under these conditions a researcher must review the units of government, typically at the national level, and determine which department would most likely control research on the topic of interest.

This review may create a list of several possible units or departments. Contact must then be made with each of these identified units and, usually, personal visits must be made to the head of each department. The reviewer must make inquiries as to the potential relevance of the unit's work to the topic at hand and must negotiate access to any potentially useful research documents. This process can be time-consuming and frustrating. However, as earlier stated, locating these government research documents can provide an immediate gauge of the level and extent of research that is likely to be available in the country.

In addition, government archives of many countries that currently provide access to the general public are now often accessible on the Internet. Some examples of these Internet sources will be provided later in this document.

## *Research handbooks and encyclopedias*

Two excellent resources for locating research and research documents are educational handbooks and encyclopedias. For decades publishers in many countries of the world have produced collections of research on various educational topics, edited and written by nationally and internationally acclaimed experts.

If a handbook on, say, reading education can be found, that means that the editors of the book and the contributing authors have been validated by the publisher as experts, and the work presented can be taken as representing, within reason, the state of the art in that discipline. The research reviewer can then feel confident in contacting those editors and contributing authors for help in locating relevant and useful research on their subject.

Additionally, the chapters written in the edited handbook can be consulted for references to research, and for critique and synthesis of the research they present. This facilitates searches for additional material. Hence, educational handbooks and specialized encyclopedias can provide an excellent starting point for locating validated experts, basic research reports, relevant journals, publishers, etc. from which extended personal contacts, library and computer-based searching can commence.

Finally, another positive aspect of handbooks and specialized encyclopedias is that the research presented is typically drawn from a broad base, both theoretically and practically, and it is less likely to be limited in the geographical contexts that it represents. Both of these characteristics increase the potential of the research to be useful for less commonly represented geographical locations.

However, there are several limitations that must be considered when using handbooks. First, handbooks are often 'slow to press'. It is not uncommon for a handbook to take two to three years to publish, sometimes rendering the material at least a few years out of date. However, the research included in handbooks is typically of such high quality that a few years' delay is not overly problematic. Additionally, if the research reviewer follows up on the works and authors cited, more recent research can typically be found by the same authors still working in the field.

A second issue to consider is that handbooks and encyclopedias are usually redone about every five to ten years. This in itself is recognition of the point that research can become dated rather quickly. A research reviewer should always make certain that the handbook or specialized encyclopedia being used is the most recent version available, if the most current research is being sought. If an historical approach is being used, the reviewer should check to see how many editions of a handbook or specialized encyclopedia have been published, and should attempt to obtain all versions for comparative purposes.

The main point, however, is that a given handbook or specialized encyclopedia being used in a research review may be ten or more years old, and thus be due for renewal in the near future. This is not a serious concern, but is an issue that should be monitored. The research reviewer should simply note the date of publication and take into account the potential impact of how recent the material is. Again, however, even older handbooks and specialized encyclopedias provide information concerning experts (editors) who can be contacted, the names and locations of researchers (contributing authors) in the field, and the names of publishers interested in the area.

Following is a list of some handbooks and specialized educational encyclopedias representing the types of themes common in educational research, as well as a list of the publishers best known for offering these reference sources. An Internet, library, or online bookseller (for example, amazon.com) search can easily uncover additional topics, themes, publishers and sources of this type.

*Handbooks.* Following is a brief list of some of the types of handbooks that are useful in educational research:

Hallinan, M.T. (Ed.). 2000. *Handbook of the sociology of education*. New York: Kluwer Academic/Plenum Publishers.

Hladczuk, J.; Eller, W. (Eds.). 1992. *International handbook of reading education*. Westport, Connecticut: Greenwood Press.

Kamil, M.L.; Mosenthal, P.B.; Pearson, P.D.; Barr, R. (Eds.). 2000. *Handbook of reading research*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Keeves, J.P. (Ed.). 1997. *Educational research methodology, and measurement: An international handbook* (2nd ed.). Cambridge, UK: Pergamon.

Sikula, J.; Buttery, T.; Guyton, E. (Eds.). 1996. *Handbook of research on teacher education* (2nd ed.). New York, New York: Simon & Schuster Macmillan.

Smart, J.C.; Tierney, W.G. (Eds.). 2000. *Higher education: Handbook of theory and research*. New York, New York: Agathon Press.

Spodek, B. (Ed.). 1993. *Handbook of research on the education of young children*. New York, New York: Macmillan Publishing Company.

Teddlie, C.; Reynolds, D. (Eds.). 2000. *The international handbook of school effectiveness research*. London: Falmer Press.

Wagner, D.A.; Venezky, R.L.; Street, B.V. (Eds.). 1999. *Literacy: An international handbook*. Cumnor Hill, Oxford: Westview Press.

*Specialized encyclopedias.* Following is a brief list of some of the types of specialized encyclopedias that are useful in educational research:

Alkin, M.C. (Ed.). 1992. *Encyclopedia of educational research* (6th ed.). New York, New York: Macmillan.

Clark, B.R.; Neave, G.R. (Eds.). 1992. *The encyclopedia of higher education*. Oxford: Pergamon.

Husén, T.; Postlethwaite, T.N. (Eds.). 1994. *The international encyclopedia of education* (2nd ed.). Oxford: Pergamon.

Plomp, T.; Ely, D.P. (Eds.). 1996. *International encyclopedia of educational technology* (2nd ed.). Oxford: Pergamon.

Postlethwaite, T.N. (Ed.). 1988. *The encyclopedia of comparative education and national systems of education*. Oxford: Pergamon.

Saha, L.J. (Ed.). 1997. *International encyclopedia of the sociology of education*. Oxford: Pergamon.

*Publishers.* While many publishers produce educational handbooks and encyclopedias, the following are well known for their focus on these materials. These four publishers are presented in alphabetical order. Publishers can be contacted directly by traditional mail, electronic mail, telephone or fax for a current and past listings of educational research handbooks and specialized encyclopedias. Often, pricing for these materials is different for educational organizations such as universities, schools, or government units. Thus it is advisable to indicate in your communication that the resources are being sought for educational institutions.

***Falmer Press***
Web site: http://www.tandf.co.uk/homepages/fphome.html
Offices:
Falmer Press
11 New Fetter Lane
London EC4P 4EE
Tel: +44 (0) 171 583 9855
Fax: +44 (0) 171 842 2298
E-mail: info@tandf.co.uk

***Kluwer Academic Press***
Web site: http://www.wkap.nl
Offices:
Kluwer Academic Publishers
P.O. Box 17
3300 AA
Dordrecht, The Netherlands
Tel: +31 78 639 2392
Fax: +31 78 639 2254

***Macmillan Publishing Company***
Web site: http://www.macmillan.com
Offices:
Macmillan Publishing Company
866 Third Avenue
New York, NY 10022

***Pergamon Press / Elsevier Science***
Offices:
Bd Langford Lane
Kidlington
Oxon, OX5 IGB, England
Tel: +44 (0) 186 584 3000
Fax: +44 (0) 186 584 3986

## *Inter-library loan*

When local libraries and government archives do not have sufficient books, documents and/or reports on the topic of interest, another library-based solution is the 'Inter-library loan' (ILL). Most libraries, particularly national and university libraries, have agreed to a system whereby many of their holdings can be loaned for some period of time to libraries in other cities, regions, or countries. In the case of periodical articles, government documents, or other materials of limited size, a photocopy of the requested item is sent instead of the original source.

### *Online Computer Library Centre (OCLC)*

Worldwide, the OCLC is the best single resource for ILL services. Through the programmes of OCLC's World Catalogue (WorldCat) system, resources no single library could possibly possess are made available, on a per-use cost basis. The OCLC Inter-library loan service enables libraries and patrons to electronically identify which libraries in the system have a specific item and then request to borrow that item from another library rapidly and efficiently. WorldCat is a bibliographic database and international union catalogue with 8,650 member libraries and nearly 700 million catalogue listings. Each year nearly 33,000 libraries from 67 countries around the world utilize the WorldCat ILL system.

The OCLC system does not have the ability to search for documents, only to request the loan of materials that are already identified. Therefore, other search mechanisms identified in this book need to be used first to obtain the bibliographic information necessary for the OCLC WorldCat loan request. The OCLC home page is found at the following URL: http://www.oclc.org/home/

Information on the use of OCLC and WorldCat services can be obtained for the following geographical regions at the specified URLs:

Europe, the Middle East, and Africa: http://www.oclc.org/europe/
Latin America: http://www.oclc.org/america_latina/
Asia and the Pacific Region: http://www.oclc.org/ap/

### *'Paper' requests for ILL*

While online ILL requests are typically more easily expedited than paper requests, many people do not have access to the necessary technology. If a computer request is not possible, some ILL libraries and services still accept 'paper' form requests. Many libraries that are accessible via the Internet also provide electronic requests for these loans. Remember, however, that a library on behalf of an individual patron makes ILL requests. Libraries will generally not loan materials to individuals without a library sponsoring or assuring the request. Therefore, the reviewer must locate a library that is able to make the loan request on his or her behalf.

This simplified example demonstrates that the essential bibliographic information, including the name of the library from which a document is being requested, must be supplied to implement a paper-based ILL request. Most ILL forms will also include information regarding costs of ILL services, penalty fees for overdue materials, and the legal obligations of the requesting patron.

## Online and Internet[1] sources

One of the most exciting innovations in research work has been the expansion of the Internet[2] to almost complete public access. However, while the Internet has vastly expanded the amount of information available to the general public, the quality of the information is usually unevaluated. Therefore, one of the main challenges in using the Internet as a source of research literature is that much of what can be found has not been reviewed by competent, established scholars in the various disciplines. Web-based information must therefore be very closely scrutinized, reviewed, and evaluated, unless it comes from a more traditional source that has made materials available on the WWW.

---

[1] The terms Internet, World Wide Web, Web, and WWW will be used interchangeably in this document.

[2] A glossary of basic terms used with the Internet is provided in Appendix 1.

The possibilities for expanded, often cost free, access to materials that previously had limited access is the great promise of this medium. The following discussion will elaborate on the more common sources available and important issues involved in entering this new medium of research and work.

### *The World Wide Web*

The World Wide Web, defined in its simplest form, is the system by which any computer with the appropriate equipment can connect and communicate with any other similarly equipped computer. Additionally, individuals or organizations can make their products or information available to anyone with the ability to connect to the Internet. Libraries, governments, universities and schools, and selected 'for-profit' companies make these products and information available. Thus, Internet sources of information are of particular interest in searching for and acquiring educational research documents and materials.

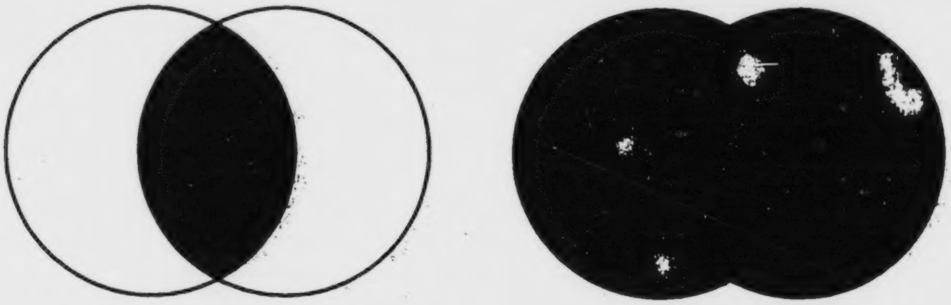
Virtually any computer that is operational can connect to the Internet if it can be connected, through a service provider, to the Web using some type of telephone system. Typically this connection is accomplished via a modem (see *Appendix 1*), although in some locations more direct and speedy connection systems are possible.

### *Combining search terms*

When performing computer searches to locate research documents, the reviewer must structure the search through a careful use of the same keywords used in doing traditional library work. Single terms (for example, keywords, author names), as used in traditional library searches, can be used individually in computer searches. However, an important advantage of computer searches over traditional library work is that on the computer these keywords can be combined in ways that make the search more specific than is possible in a library card catalogue.

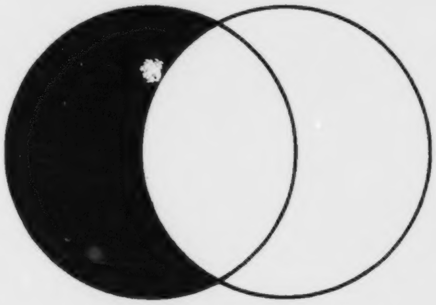
The primary mechanism for creating more specific searches on the computer is the use of Boolean operands. Boolean operands, commonly used in computer programming, are used in computer searches to specify certain types of relationships between the terms that are being used to locate research documents on specific topics.

**Figure 2.3 The use of operands in searches**



Education **AND** Ukraine

Education **OR** Ukraine



Education **NOT** Ukraine

As shown in *Figure 2.3*, the most basic application of Boolean operands is to combine two keywords. In *Figure 2.3* the result of using the same two terms with each of the three primary Boolean operands is graphically displayed. If the 'universe' of potential research documents is represented by the contents of a circle, then there are two potential 'universes' of research documents being searched in *Figure 2.3*: (a) all potential research documents that focus on education, and (b) all potential research documents that focus on the Ukraine. Clearly, not all research in education will be appropriate to

BEST COPY AVAILABL

this specific research review topic, and not all research on the Ukraine will deal with education. However, for the purposes of this example these two terms will be used.

In the first example in *Figure 2.3*, if the Boolean operand **AND** is used to designate the relationship between education and the Ukraine, the computer search will identify only those items that include both the terms *education* **AND** *Ukraine*. In mathematics, this would be called the *intersection*. This search would exclude all items that include education but not Ukraine, as well as all items that deal with the Ukraine but not education. This search would be the most specific of the three and thus would yield the fewest items.

If the results of using the Boolean operand **AND** did not yield sufficient material, a broader search would be necessary. Using the second operand **OR** would produce the broadest search, yielding the largest number of items of the three Boolean terms being considered. In mathematical terms this would be the *union*. If the **OR** operand is used to connect the two terms in a computer search, the result will identify all items that utilize the term *education* (one of the most common terms in computer databases), whether or not Ukraine was involved in the research, as well as all items that deal with the Ukraine, regardless of whether education is involved. Clearly, this search would produce numerous items that would not interest someone doing a research review on education in the Ukraine. But this very general search would capture the broadest set of items possible.

Finally, the third operand, **NOT**, would be used if the interest of the reviewer focused on any educational research literature except those items dealing with the Ukraine. This is probably the least used of the three basic Boolean operands, but it can be important to understand when the use of the two more common terms is not producing the desired results.

This presentation has focused on the use of operands in the most basic form: to establish a specific relationship between only two terms. Most computer searches will allow the use of compound (nested) terms, such as the following:

((Education **AND** Ukraine) **NOT** Tertiary)

In this nested search, the computer will search only for those items that include both of the keywords, *education* **AND** *Ukraine*, but will exclude any of those qualifying items that also focus on the tertiary system. The use of parentheses in a nested search indicates to the computer exactly which terms are to be combined by a specific operand. Thus items dealing, most likely, with primary and secondary education in the Ukraine will be retrieved. The careful and creative use of nested search designs can yield highly specific results and avoid the necessity of reviewing numerous items that are not even closely related to the topic of the review.

### *Subscription services*

Many Internet-based resources that contain information on research projects require that an institution such as a university or library pay a fee, typically annually, for access to their base of information (database). These annual fees are often accompanied by a surcharge for each report or article copied from the database. Sometimes a fee for the time that a researcher is actually connected to the service's database is also levied. Thus subscription services can be costly, and before too much work is done on these services a careful review of the potential costs and benefits should be made.

Historically, these services were usually the best and most comprehensive compilations of research documents. In recent years, however, non-subscription services have increased in both numbers and comprehensive services. Some previously fee-based systems have even converted to become almost entirely free, such as the Educational Resources Information Center (ERIC) based in the USA (http://www.edrs.com/).

Since the usefulness, or at least the necessity, of subscription services has largely been undermined by free services in recent years, a discussion of the remaining providers will not be presented here. If these services are of interest, however, a review of several sources referencing them could be made.

*Free services*

There are numerous Internet sites where statistics and research related to educational issues can be located and usually downloaded to your computer without charge. If the documents cannot be downloaded without charge often the paper version of a report can be obtained at little or no cost. Often books and reports can be ordered. Additionally, most sites include links to related locations on the WWW. The main things to keep in mind when attempting to access these free services are to be patient, to be willing to explore the site, and to *keep notes* on which sites have been searched and what the results of those searches have been. A few sample Internet sites are provided in the following sections, and others are listed in *Appendix 2*.

*The European Union (EU).* Documents provided by the EU related to education and statistics are available at the following sites:
http://europa.eu.int/pol/educ/index_en.htm
http://europa.eu.int/
http://www.eurydice.org

The EU Internet site is accessible in 11 languages, but not all documents are available in all of them. Most documents are available in English, and many are in French and German.

*Organisation for Economic Co-operation and Development (OECD).* The data, texts and products of the OECD can be searched using the home page as a starting point: http://www.oecd.org/

The texts on these sites are mostly in English, and some are available in French. Many documents on these sites are in PDF file format, which requires the software package *Acrobat Reader. Acrobat Reader*, along with instructions for downloading and using the software, can be retrieved on the Internet free of charge at the following web address (URL): http://www.adobe.com.

The primary search program for the OECD is available at the following URL: http://www.oecd.org/search/

*United Nations Educational, Scientific, and Cultural Organization (UNESCO).* At the present time the UNESCO home

page is available in English, French, Spanish, Russian, and Chinese. However, once past the home page most information is in English, and some is available in French. Most documents on the UNESCO sites are available in PDF format, and many sites provide the opportunity to download the *Acrobat Reader* directly from the site. The UNESCO home site can be located at the following URL:
http://www.unesco.org/

UNESCO's World Education Forum is available at the following URL:
http://www2.unesco.org/wef/

UNESCO documents can be searched and downloaded (typically in PDF format) or ordered at the following sites:
http://www.unesco.org/education/
http://www.unesco.org/education/unesearch.html
http://www.ibe.unesco.org
http://unescostat.unesco.org/ (this site includes data on national education systems)
http://unesdoc.unesco.org/ulis/
http://unesdoc.unesco.org/ulis/ged.html

The UNESCO Library is available at the following URL:
http://www.unesco.org/general/eng/infoserv/doc/library.html

The UNESCO Library site allows access to the UNESBIB (bibliographic records of UNESCO documents, publications, and library collections). And the UNESCO Thesaurus, which provides information regarding the most likely terms to use in searching the UNESCO site, is available at the following URL:
http://www.ulcc.ac.uk/unesco/

The site of the International Institute for Educational Planning/ UNESCO also provides access to specialized bibliographic records and publications:
http://www.unesco.org/iiep/

*Australian Council for Educational Research (ACER).* The documents and resources of ACER are available at the following sites:
http://www.acer.edu.au/
http://www.acer.edu.au/acer/research/search1.html.

The ACER site allows you to do keyword searches. From the main ACER site you can access Request for Documents. The site gives you directions and fees for ordering library material (journal articles, conference papers, and book chapters). Books and reports can be requested by the interlibrary loan system. From the library site you get information for accessing the Australian Education Index (AEI) database via the Internet through Informit Online or via CD-ROMS. While much of ACER is a free service, extended services and access is a system requiring a fee. At the time of writing, the price for the Informit Online subscription is US $400 for a single user and US $200 for an additional user. Basic and secondary schools receive discounts. Users in developing countries can also get a subscription discount on the price.

*United Nations Children's Fund (UNICEF).* All UNICEF sites are available in English, and some are provided in French and Spanish. The main home page of UNICEF is accessed at the following URL: http://www.unicef.org/

UNICEF documents and publications, including the UNICEF Annual Reports, which are in PDF format, are available at the following URL: http://www.unicef.org/apublic/

Statistics produced and kept by UNICEF, which can be accessed by country, are available at the following URL:
http://www.unicef.org/statis/

The UNICEF Research and Evaluation materials are available at the following URL: http://www.unicef.org/reseval/

*Educational electronic journals.* The 'Communication of Educational Research' special interest group (SIG) of the American Educational Research Association (AERA) provides this site, which provides access to nearly 70 electronic scholarly journals that are full text and free of charge. Some journals offer a graphic version for fast Internet connections and a text version for slower connections, such as modems. Articles from these journals, available in HTML or PDF format, and available at the following URL:
http://aera-cr.ed.asu.edu/links.html

On their Internet home page (see URL immediately preceding) this site is described as including

"...only links to electronic journals that are scholarly, peer-reviewed, full text and accessible without cost. We have excluded professional magazines that are largely not refereed, and commercial journals that may only allow access to a very limited number of articles as an enticement to buy. By restricting membership in this way on the list that follows, we hope to do what little we can to promote free access worldwide to scholarship in education..."

*World Bank*. The web sites for the World Bank can all be searched by keyword. The principal World Bank home page is located at the following URL: http://www.worldbank.org/

From that location World Bank educational research and material can be accessed through 'clicking' on the appropriate icons, or by going directly to the following sites:

http://www.worldbank.org/research/
http://www.worldbank.org/data/
http://www.worldbank.org/html/schools/data.htm
http://www.worldbank.org/html/schools/research.htm

World Bank sites provide access to their Digital Library, Journals, and Policy Research Working Papers.

*University libraries online*

The library holdings of most major universities throughout the world are available for searching via the Internet. Once at the home page for a university, there is usually an icon that can be activated for the library. The holdings of these libraries can be searched for books and documents, but the full texts of those items are not available online, largely due to copyright issues. However, some libraries are beginning to make their materials available in digital formats that can be downloaded. But in most cases the library materials of specific universities located by searching on the WWW need to be obtained through the inter-library loan system. A few Internet locations for well-known universities are listed below, as well as one site that allows access to a consortium of university libraries.

*Gabriel.* This gateway is a particularly valuable resource and provides access to the 40 national librairies in Europe. Each national library linked through this site provides its own description of the system for procuring documents. *Gabriel* is situated at the following URL: http://www.portico.bl.uk/gabriel/

*Consortium of English Libraries (COPAL).* COPAL provides access to the library holdings of 19 well-known libraries in the United Kingdom, including Oxford, Cambridge, Edinburgh, Glasgow, and the University of London, among others. The texts at these libraries are available in over 100 languages. On the Web, COPAL is located at the following URL: http://copac.ac.uk/copac

At the moment it is not possible to borrow a book or journal article directly via COPAL. An Inter-library loan request must be made through one's local academic library. The COPAL document information is the basis for the Inter-library loan request.

A photocopy of a journal article or other text can be obtained by Inter-library loan, or a personal request for article photocopies can be made to the British Library using their Articles Direct service. To access this service on the British Library Web page (http://www.bl.uk), select the Articles Direct icon for details and instructions.

*Harvard University.* The library holdings of Harvard University can be searched on the Harvard OnLine Library Information System (HOLLIS). The Harvard library system is the oldest in the USA and the largest academic library in the world. Books, manuscripts, microforms, maps, slides, photographs, and other materials are housed in more than 90 individual collections. To search the 8.5 million indexed items of the larger library holdings, which are distributed throughout all 90 locations, the Harvard Union catalogue (HU) can be selected at the following URL: http://hollisweb.harvard.edu/

*Université de Paris-Sorbonne.* This web site, in French, provides access to many educational documents, research reports and document services. The library holdings of the Sorbonne University in Paris, as well as the 40 national libraries in Europe (via the *Gabriel* gateway) and 96 university libraries throughout France (via the *les*

42

*bibliothèques universitaires françaises* hotlink) can be accessed at the following URL:

http://www.paris4.sorbonne.fr/html/biblio/index.htm

*University of Peking (Beijing).* The web site for the University of Peking provides both Chinese and English language interfaces to the general resources of the university library. However, the most detailed access and holdings are in Chinese. The University of Peking is currently experimenting with a digital library system (supported by the New Zealand digital library program), as well as other electronic service systems. However, these are still limited in scope. The home page for the university is located at the following URL:

http://www.lib.pku.edu.cn/

This web site provides several important links. The most prominent is the *China Academic Library & Information System* (CALIS). Offered in Chinese, the CALIS system provides extended access to documentary and research material throughout China.

# III. Analyzing quantitative research

Once potentially useful research documents have been acquired, the reviewer must undertake a careful, systematic analysis of both the quality and potential usefulness of the findings. This analysis should evaluate the contextual relevance of the sample and data to the specific policy setting of interest, as well as the technical merits of the research. Thus, each research document should undergo careful analysis based on both internal quality issues and external issues.

## *Reliability and validity*

The central aspects of the scientific usefulness of any quantitative research work are reliability and validity. Unless the reliability and validity of a research project are considered to be acceptable, the results of the project are suspect and perhaps unusable. Thus the first and ultimately most important two questions that must be addressed in assessing the quality of any single research document are: (a) does the article address and present evidence of the reliability of the instruments used and data collected, and (b) is there clear indication regarding the validity of the study being evaluated?

### *Reliability*

Traditionally, reliability is considered to be the ability of the instrument used in research to consistently measure the characteristic of interest. An instrument is considered reliable to the degree that it consistently measures the characteristic of interest (i.e. academic performance, skills acquisition, etc.), both over time and from subject to subject (Gay and Airasian, 2000). The more reliable an instrument is, the more confidence can be placed in the results of the measurement. For example, if the mathematics scores of students in Krakow, Poland, were measured with a certain instrument in September 2000 and then were measured again six months later, with the same instrument producing reasonably the same results, then the mathematics ability instrument and the results would be considered reliable. However, if the attitudes of a group of students towards education in Ottowa, Canada, were measured on two different occasions, and the rank order of students on the attitude scale was quite different on each occasion, then the stability of either set of scores would be in serious doubt, and hence considered unreliable.

Obviously, reasonably high levels of reliability are preferable in research that is used in a review. If there is serious question as to the reliability of data from the research as originally conducted, then the potential usefulness of the results in another place and/or time is significantly diminished. Many things can and do influence reliability and, to a significant degree, these potential threats to reliability are what a research reviewer is attempting to access as each research document is evaluated.

Often the reliability of data can be influenced by the way questions are formulated and asked. Even seemingly direct demographic information, such as age, family configuration, etc., can be rendered unreliable by the way that the questions are designed and delivered. Usually the reliability of this type of data remains untested, which can negatively influence the usefulness of the resulting data. *Table 3.1* presents a few examples of the reliability of data being influenced by how a question is asked.

**Table 3.1 Examples of more and less reliable questions for a survey**

| More reliable | Less reliable |
|---|---|
| Q. How old were you, in years, on your last birthday? | Q. How old are you? |
| Q. In what city and nation were you born? | Q. Where were you born? |
| Q. How many full biological brothers and sisters do you have? | Q. How large is your family? |
| Q. How many years of formal education have you completed? | Q. How much education do you have? |

As each research document is evaluated, the reliability of the data must be assessed. Sometimes statistical information is provided regarding the technical qualities of the reliability of the instruments used. However, more often the assessment of reliability is left to the reviewer's professional critique, experience, common sense and, frankly, intuition. There are several areas where data can be provided for evaluating the reliability of an instrument and the data it produces: stability, internal consistency and inter-rater reliability.

*Stability*. This is the type of reliability that was discussed above. Stability is the degree to which the same test or instrument yields consistent results over time. The most common indication of stability in a research document would be a statistical value (a correlation coefficient indicating test-retest reliability) relating the 'coefficient of stability'. This would be presented as a number between 0 and 1.0. A coefficient of 1.0 would indicate perfect (100 per cent) reliability, while a coefficient of 0 would indicate no reliability (0 per cent). Obviously numerical values between 0 and 1.0 indicate weaker and stronger reliability in the instrument and, therefore, in the results.

There is no clear level at which the stability of an instrument is deemed to be acceptable in any single case, let alone in every case. In certain cases, an instrument being used in an attempt to measure a complex variable such as 'attitude' with a lower coefficient (0.27, for example) might be considered more acceptable than an instrument with a higher coefficient (such as 0.34) used to measure a more simple and direct variable such as arithmetic performance. The evaluation of stability is an exercise of common sense and experience, of balancing the relative complexity of the variable being measured and the likelihood of a higher or lower coefficient of stability.

If no indication of stability is given, then the results of the research should be used with caution. Even low coefficient values are better than no information at all.

*Internal consistency*. Internal consistency is a common form of reliability that provides information about the consistency of items within a single instrument or test. The most common indication of internal consistency is a measurement of the split-half reliability of an instrument.

The procedure for this approach is to administer the instrument or test to a group. Then the test is divided into two comparable halves, usually done by grouping the even and odd-numbered items. The two newly 'created' sub-tests, which are assumed to be roughly equivalent, are then tested for correlation between the scores using the Spearman-Brown correction formula. This produces a score between 0 and 1.0. As discussed previously, scores closer to 1.0 are considered better in terms of internal consistency, and scores closer to 0 are considered to indicate an unreliable instrument.

*Inter- and intra-rater reliability.* When an instrument contains elements that are subjectively scored, such as essays, short-answer tests, performance and product tests, projective tests, and/or observations, the reliability of the instrument and the data produced are sometimes called into question. In such conditions the concern is for inter-rater (inter-scorer, inter-observer, etc.) reliability and/or intra-rater reliability. Inter-rater reliability refers to the consistency between two (or more) independent scorers. Intra-rater reliability refers to the consistency of a single scorer over time.

Estimates of inter-rater or intra-rater reliability are usually provided either by use of correlational techniques (similar to those discussed in the previous sections on stability and internal consistency) or as a simple percentage of agreement. Inter-rater reliability can easily be tested by having two (or more) raters score the same instrument independently and then comparing the scores produced either by correlation or by percentage of agreement. Intra-rater reliability is tested by having a rater score the same test at two different times and then comparing the two scores.

The more open-ended the scoring is, the more important it is to establish inter- and intra-rater reliability. Subjective scoring always reduces reliability and, therefore, diminishes the ultimate validity of the interpretations, conclusions and generalizations of the research.

### *Validity*

Although reliability is important to the usefulness of a research document, validity is much more critical. Validity is critical in all forms of research and in all types of tests and measures. Since research that is being reviewed comes, potentially, not only from different geographical areas but also from varying times in the past, validity must be considered both as it was viewed historically and as it is construed at present.

*Historical notions of validity.* For many years, in some educational disciplines as recently as the past decade, validity was viewed as a characteristic of the instrument or test itself. An instrument was seen to 'possess' validity, independent of how or why it was used. In this sense, validity was assumed to consist of three different types: (a) content validity, (b) criterion-related validity, and (c) construct validity.

(a) Content validity refers to the extent to which a test measures a representative sample of subject-matter knowledge and behavioural skill from the course of study that is to be assessed. For example, consider a test that was designed to measure competence in using the English language. In order to examine the content validity of this test it would be necessary first to examine the subject-matter knowledge and behavioural skills that were required to respond correctly to the items on the test. These would then be compared to the subject-matter knowledge and behavioural skills that were widely accepted as comprising the correct and effective use of the English language. If the comparison showed a large overlap then the test could be described as having high content validity.

(b) Criterion-related validity refers to the capacity of test scores to predict future performance, or to estimate current performance on some valued measure other than the test itself. For example, a reading readiness test might be used to predict future reading achievement, or a test of dictionary skills might be used to estimate the capacity to use a dictionary (as determined by observation). In the first example, we are interested in prediction (the relationship between two measures over an extended period of time). In the second example we are interested in estimating present status (the relationship between two measures obtained concurrently).

(c) Construct validity is concerned with the extent to which test performance can be interpreted in terms of psychological constructs. A construct is a psychological quality which we assume exists in order to explain some aspect of human behaviour. For example, 'intelligence' is a construct. When we interpret IQ test scores as measures of intelligence, we are implying that there is a recognizable quality associated with individuals that can be properly described as 'intelligence' and that this quality can account to some degree for performance on the test.

If a research document that was produced more than 10 or 15 years ago is being evaluated for use in a research review, the approach taken to establishing the validity of the instrument, and hence the research itself, is likely to draw upon one or more of these three types of validity. Evidence should be sought regarding all of them, but convincing evidence regarding construct validity is particularly important because construct validity indicates to what degree the instrument is really measuring what it claims to measure. When a research document provides convincing evidence that the construct validity of the instruments used is high, then the results of the measurement are potentially the most useful.

Obviously, finding evidence in all three areas of validity would be preferable, especially if all three areas are persuasively presented as being strong. In practice, however, few research documents provide information in all three areas, and many provide no indication of the validity of the instruments. If no evidence is provided, a reviewer has no choice but to consider the results tentatively valid, at best. When evidence is provided, the hierarchy of preference is as follows: first, construct; second, criterion; and third, content validity.

*Contemporary notions of validity.* Even though validity is being presented here as both historical and contemporary in use, it should not be construed that construct, criterion, and content validity are no longer used at all. This is far from the truth. However, they are used in a much more holistic fashion at present than they may have been in the past.

Whereas in the past validity was seen as a quality of the test or instrument itself, now validity is taken to be the result of the interaction of the test, the test administrator, the conditions under which the instrument is delivered, and the ways the results of the instrument are applied. It is no longer simply stated 'the test is valid'; rather, it is now said that 'the test is valid for this particular interpretation, this specific use, and this particular group'.

Clearly validity is a crucial feature of a test or instrument, and in this respect construct, criterion, and content validity are still applicable and vital in determining this quality. Even a test considered to be highly valid, according to these three aspects of validity, can be misused and over-interpreted, however, in which case the research would be rendered unreliable and invalid to varying degrees.

An additional dimension of validity that has been recently added is the concern over the consequences that arise from the use of tests and instruments. The social and cultural impact of the use of even a highly 'technically' valid instrument must be considered when making contemporary determinations of validity. The use of an instrument that would probably produce negative social consequences (for example, the placement of students from a traditionally vulnerable social group or class into educational situations with limited resources and lower access to skills and job training) would not necessarily be considered a valid use of a test in the present global context.

At present, then, when considering the validity of a specific research document or group of research documents, the reviewer must account not only for traditional notions of instrument validity, but also for a more holistic concept of the use, interpretation, and potential social and cultural impact of the use of data derived from a particular test. This task is not a technical one; a correlation coefficient cannot be calculated for the interpretational validity of research. This more holistic application of validity is largely a function of the reviewer identifying, carefully considering, and then drawing conclusions concerning these extended contemporary elements and notions of validity. In this regard, then, validity has gone beyond what can be rendered quantifiably, even though some aspects of it (i.e., test-based construct and criterion validity) still are.

*Concluding thoughts on reliability and validity.* Although considerations of reliability and validity are the cornerstones of the potential usefulness of research that is being evaluated and reviewed, unfortunately very few reports and articles provide sufficient information in these areas. Thus to a large degree the most important and powerful elements of review and evaluation are typically limited or unavailable.

But this unavoidable challenge to assessing reliability and validity should not prevent the review and use of research that is difficult to assess. Rather, after a conscientious effort has been made to locate evidence for the reliability and validity of the research, the reviewer should persevere, maintaining a healthy concern for the potential limitations of the research.

Insufficient evidence of the reliability and validity of the research should heighten the intensity of review on the aspects of the research document that are more readily described. Scrutiny of subsequent elements should then be expanded, and the centrality of each subsequent element increased. The task of locating dependable and credible research for review is not made impossible by the lack of evidence for reliability and validity; it is simply refocused to a more careful consideration of certain elements, which will be discussed in the following chapter.

# IV. Components of research documents for purposes of analysis

Most research documents contain materials that need to be evaluated to determine their usefulness for a review. Ideally, information from research documents will address each of the eight general areas of the research review (mentioned in *Chapter 1*). However, a more realistic expectation would be for sufficient information in some of these areas to indicate general usefulness of the research results. As very few research documents will provide all of the information preferred, a reviewer should approach the task of reviewing with a realistic and flexible attitude. Never discard a report because it provides little detailed information for evaluation. Rather, consider the potential threats to the usefulness of the results, and move ahead with the review.

Each research document should be scrutinized to establish two things: the internal validity of the research and the external validity of the research. Internal validity refers to whether the research has been done well enough to be valid for the population intended, and external validity refers to whether the research has validity to populations external to the original. It is important to distinguish between these two types of validity because research might be very well done and therefore be applicable to the intended population (high internal validity) but not be useful at all for any other population (low external validity). It is easy to assume that good research is applicable anywhere in the world with any population – but this is not true.

The following section addresses each of the eight research review elements in terms of more specific evaluation issues and criteria. Worksheet 4.1 is provided to guide the evaluation of each research document as it is being reviewed.

## Worksheet 4.1 Checklist for reviewing research documents

---

Mark (X) *when complete*

**Purpose and problem(s)**

- Purpose found and noted? ☐
- Problem(s) found and noted? ☐
- Variables (educational conditions) of interest specified? ☐
- Context of research clarified? (educational settings, sociocultural issues, etc.) ☐

**Research method and methodology**

- Methods (e.g. survey, interview, archive) presented and discussed? ☐
- Methodology described? ☐
  - Descriptive? ☐
  - Correlational? ☐
  - Causal? ☐
- Methods and methodology appropriate to each purpose and problem? ☐

**General procedures**

- Discussion of procedures sufficiently detailed? ☐
- Procedures appropriate and reasonable for context of study? ☐

**Sampling**

- Population defined? ☐
- Sample selection process specified (randomized)? ☐

**Measurement**

- Instrumentation described? ☐
  - Reliability discussed? ☐
  - Validity discussed? ☐
- Use of instruments reasonable for context? ☐
- Training of researchers discussed and adequate? ☐

**Data analysis**

- Analysis appropriate for purpose, problem(s), and methodology? ☐

Statistical information complete? ☐
Data used in appropriate and scientific way? ☐
Statistical and functional significance considered? ☐

**Results**

Purpose and each problem addressed completely and directly? ☐
Problems or issues presented that were not in original design? ☐

**Conclusions and generalizability**

Purpose addressed completely and directly? ☐
Each problem addressed completely and directly? ☐
Internal validity high? ☐
External validity high? ☐
Conclusions adequately warranted? ☐

## *Purpose and problem*

The first element of the research review process is identifying the purpose and problem(s) stated for the research. The purpose is the reason for conducting the research. Purposes range from the lending requirements of international funding agencies, to government programme design and implementation needs, to degree (thesis and/or dissertation) requirements for university graduation, among many others.

Understanding the purpose is critical because each of the subsequent seven elements is narrowed and designed on the basis of the purpose. Once the purpose is identified, certain problems become more or less reasonable to pursue.

Given a stated purpose, a research project typically identifies a specific problem or set of problems to examine. A problem statement indicates the variables (for example, educational conditions, student/parent/teacher demographic characteristics) of interest to the researcher, along with the specific relationship between those variables (Gay and Airasian, 2000; Krathwohl, 1997). However, in reality, most research can only explore a small 'piece' of the identified problem area, and this is selected in terms of the purpose of the project. This narrowing of focus may create several challenges to the use of research, and these will be addressed in later sections of this book.

If the research document is well written, both the purpose for conducting the research and the specific problem(s) derived from that purpose will be explicitly stated.

When this is so, then the reviewer's task is much more simple, as the rest of the research document can be reviewed for coherence with and support for these important guiding elements. For example, the results and conclusions section at the end of the research document should relate directly to the purpose and problem(s), providing data and evidence for each one. If the results and conclusions section 'skips' a presentation of how the purpose was met or how any of the stated problem(s) answered, then the research document is deficient in terms of internal coherence.

Hopefully, the sampling, data analysis, etc. presented in the research will both clearly reflect on the purpose and problem(s) and represent solid methodological procedures. If so, then the likelihood increases that the internal validity of the research is high. In contrast, if the sampling or other design elements do not coalesce with the function and intent of the study, then the internal validity of the research would decrease with each infraction. Establishing that the research is internally sound is a prerequisite to establishing its external value, and therefore to the functionality and potential usefulness of the research review. A study cannot be weak internally and still be used with credibility in another setting.

If either the purpose or problem(s) are not stated explicitly, then the reviewer may conclude that data collection and analysis were not guided by a distinct purpose or question, or may attempt to derive a purpose and problem from the information presented. Deriving or inferring the purpose can be dangerous, especially when the setting of the research is distant in terms of space, time, or cultural context from the reviewer. It is typically advisable to simply note the lack of clear indication of purpose or problem and then move on to the other review elements. The misattribution of purpose and problem can lead to extremely negative consequences and should be avoided with reasonable judgement.

The most challenging task of the reviewer is not in establishing the internal validity of the research, although this is indispensable, but rather in determining to what degree the purpose and problem(s) are applicable to the potentially different condition and setting of the review. Consider, for example, that a specific research document may have examined a condition where the local governing agency (for example, school board or parent group) had required the evaluation of the teaching experience of the secondary-school staff in a specific town and assessed the resulting impact on student performance. If the review of the procedures of the research document produces evidence that the research was well done methodologically and procedurally, remaining focused on the purpose and problems, then a conclusion of good internal validity might be reached.

If the subject of the research review was also the influence of teacher experience and student performance, since the internal validity was judged to be good the reviewer might easily assume that the apparent match in subject combined with solid internal validity automatically makes the study externally valid for the review being constructed. This could easily be an inappropriate conclusion. If, for example, the research being reviewed was conducted in a developing country where teacher experience was not influenced or augmented by continuing teacher education at tertiary institutions, and the location of the research review was in a developed country where access to such in-service professional development was quite common, then the results might not be usable due to the contextual differences. The potential confounding influence of these sorts of context-specific intervening variables, or lack thereof, are critical in evaluating the usefulness of research to settings other than the original context (external validity).

The reviewer must remember that multiple research documents with similarly stated purposes and problems should not necessarily be treated as the same, or even as reasonably similar. When multiple research documents use the same or similar words in describing methods and procedures, it is easy to assume that they focus on the same real issues. This is risky unless it can be determined that the influence of the purposes (financial, functional, political, academic, etc.) and the setting (Thailand, Samoa, Ireland, etc.) are such that true comparison is possible. The context of the original setting must be carefully compared to that intended in the review and the differences must be taken into consideration in the review process.

## *Research methods and methodologies*

The second element of reviewing research is to consider carefully that different quantitative research methods and methodologies have strengths and weaknesses relative to the particular types of problems that can be addressed, the answers that are possible, and the conclusions that may be reached with confidence (Scott and Usher, 1999). Often research projects that are otherwise well designed overreach the limits of their basic methodologies and come to indefensible conclusions. This challenge typically occurs in policy formation when the researcher is not able to reach 'cause and effect' conclusions given the true nature of data and research available. The lack of a rational basis for causal conclusions creates challenges, as most educators, policy-makers and governments want to operate with confidence that what they do will reliably produce the desired result.

Therefore, research reviews must carefully evaluate the actual range of defensible conclusions based on the constraints and design parameters of specific methodologies. Research reviews are often seriously flawed by failure to adequately consider the limits of methodology in terms of the often professionally and politically charged conclusions that may result.

Research methods are the functional approaches taken in the research project. The same problem could be approached utilizing a number of different methods or combinations of methods. For example, a study of the effectiveness of a reading instruction programme in Thailand could be undertaken by any of the following methods, alone or in combination:

1. Survey method: survey teachers using the reading programme of interest to explore how they rate its effectiveness;

2. Interview method: interview teachers, students and/or parents to investigate the impact of the reading programme;

3. Observational method: have trained researchers observe and evaluate implementation of the reading programme;

4. Archival method: analyze the archival records of student reading performance (grades, class test scores, standardized reading scores etc. could be used to discern effects of the reading programme).

Obviously, any of these methods could be used to investigate the reading instruction programme in Thailand. No method is automatically better than another; each has its own criteria of implementation that would lead to more reliable and valid or less reliable and valid results (Gay and Airasian, 2000; Krathwohl, 1997; Scott and Usher, 1999).

While the method is the functional approach taken to the research project, the methodology is the larger context in which particular methods are applied. While there are many methodologies that are commonly referred to in the general research literature, in this treatment only three very general methodologies will be introduced here and discussed in greater detail later: (a) descriptive, (b) correlational, and (c) causal.

Descriptive methodology applies to research conducted solely for the purpose of describing the current state of an educational context or situation. It does not intend to imply any relationship between the variables or conditions that it explores. Descriptive methodology research may be the most common type of research because it uses much data and information that are either already available or can be obtained with a minimum of cost and effort.

Correlational methodology reaches to the next level of inferential complexity. It guides research that explores basic relationships between variables. Not only are variables in education described, but the association and relationships between them are explored statistically.

While correlation research cannot establish causality, correlation lays a foundation before more costly and long-term causal research should be undertaken. It is common to assert that when variables are correlated they *may* be causally linked, but if variables are not even correlated they *cannot* be causally linked.

Causal methodology is the most complex and assertive of the three general types of methodologies. Well beyond description and correlation, causal methodology focuses on establishing the causal relationships and sequences of conditions in educational settings. Causal methodology is considered the pinnacle of scientific educational inquiry, as it provides the most reliable basis for establishing educational policy and practice. However, high-quality causal research is difficult to locate, and therefore its use is rare in educational policy endeavours.

These three types of methodology are described in more detail in *Chapter 5*.

## *General procedures*

The third element for reviewing a research report is to identify the general procedures that were used. How a research project is actually conducted can influence the reliability and validity of the results, as well as the analysis and conclusions. Ultimately, the generalizability, and therefore usefulness, of the research report can be compromised by how a study was implemented.

Thus a careful evaluation of the general procedures must always be made in determining the appropriateness of a study. While the finished review of research seldom presents the details of this evaluation, it must take place before conclusions from that research are presented in the review.

The most fundamental aspects of how a research project is actually conducted comprise the procedures. If, for example, a research project is based on mailed surveys, once the survey instrument had been developed the procedures would include the following:

1. When the survey was first mailed to the subjects;
2. When the first follow-up reminder was sent;
3. How the returned surveys were registered and tracked;
4. How the data from the surveys were actually coded;
5. How the data were entered for analysis;
6. Etc.

The development of the survey instrument would also entail procedures to follow. Each procedure could be expressed in terms of the actual steps necessary to complete a task. The procedures, then, are the steps taken to implement the methods within the methodology of the research project. Not all procedures are equally valid or of equal quality.

As the reviewer evaluates each research document, the general procedures need to be carefully considered. Significant background information may be needed to accurately determine whether the procedures were appropriate for the research under review.

For example, educational research often includes a mailed survey. If a survey were first mailed to teachers under consideration in a study of teacher qualifications during the first week of June, and then a follow-up reminder were sent two weeks later, the resulting low response rate from teachers might be a mystery to the research reviewer. If the reviewer understood, however, that the regular school year for the teachers involved in the study ended in the final week of May and did not resume until the final week of August, then the fact that the survey was sent during the time that teachers were not at school would indicate inappropriateness of the survey procedure. This type of context-specific knowledge is important when reviewing educational research procedures, but can be difficult to obtain if many research documents are being reviewed from diverse areas of the world.

Hopefully, the research documents themselves will explain the limitations of their procedures and methods. However, if no such discussion is contained in the research document then reviewers must rely on their own knowledge of the context of the research to detect such subtle challenges to the research design. When a research study is evaluated in a context unfamiliar to the reviewer, then the reviewer should either seek the input of a person experienced in the setting of the research or utilize the results of the research with great caution.

## *Sampling*

The fourth element of the research review process involves understanding and accounting for the sampling procedures. Most research utilizes only a sampling of the population of interest, inferring the results of the sample to be applicable to the larger population from which that sample was drawn. But many samples deviate from the rules of scientific probability sampling and therefore provide biased estimates that cannot be evaluated with respect to their accuracy. The reviewer must therefore take a very close look at the sampling procedures to establish the potential for bias and to check that important estimates of population characteristics are provided with estimates of sampling errors.

Determining how confidently one can accept the results of any study or group of studies is difficult and complex. However some general considerations can make the task easier to organize and thus help in coming to a defensible conclusion regarding both the usefulness of research conducted with a sample for its original population and its usefulness for another group.

Sampling is the process of selecting a number of individuals (or 'units' such as students, classrooms, schools, districts, etc.) for inclusion in a research study in such a way that estimates of the characteristics of the larger group (population) from which they are chosen have no (or minimal) bias and have known confidence limits based on correctly calculated sampling errors. Sampling is commonly used in educational research because the larger population of interest in the study would be too difficult, time consuming, or expensive to study in its entirety. Hence, some smaller number (sample) of the population is used as an adequate representation for the purposes of the research. If scientific probability sampling procedures are employed, then the results of the survey can be used to make informed inferences about the characteristics of the population.

The first step in sampling is to define the population of interest, the group to which the results of the survey are intended to apply. For example, a researcher might be interested in knowing the reading ability of all 12-year-old students in Indonesia. The population of interest, then, would be all 12-year-old students in Indonesia. Clearly, there are many Indonesian students of this age, and to test each one of them would require large amounts of funding, political support, and a considerable period of time. By carefully selecting a sample of these students the researcher could complete the study in a timely and dependable fashion, and then, if the sampling was appropriately conducted, the results could be generalized to the entire population of Indonesian 12-year-old students.

Regardless of the particular sampling technique used, the procedures should be described in the research document. Each sampling technique has benefits and deficits in terms of the particular research setting. There is no one sampling technique that can be considered 'best' for all studies and circumstances. The 'fit' of a sampling technique to a particular setting is a judgement on the part of the researcher, and it should be considered carefully by the reviewer.

All sampling techniques proceed according to the same basic steps: (a) identify and define the population of interest, (b) determine the required sample size (based on sampling error requirements), and (c) select the sample according to scientific probability sampling procedures (Gay and Airasian, 2000, p. 123).

There are a number of basic sampling techniques or procedures. All of the techniques described below can be referred to as *scientific probability sampling techniques* because each element in the target population has a known and non-zero probability of selection. It is only from these types of sample design that it is possible to assess the accuracy of sample estimates. The five basic probability sampling designs are as follows:

1. Simple Random Sampling, the process by which the sample of a population is selected randomly. Each member of the population has an equal, independent, and non-zero chance of being selected

for inclusion. As the researcher does not make the selection, there is no possibility of researcher selection bias.

2. Stratified Sampling, the process of randomly and independently selecting samples within important strata or sub-groups of the target population (for example, those based on gender, race, ethnicity, tribality, academic risk categories). In the special case where the samples selected within each stratum of the target population are proportional to the total size of each stratum, we refer to the final combined sample as a stratified proportionate sample.

3. Systematic Sampling, a variation on simple random sampling in that the members of the target population are listed and the selection of the sample is made by taking every 'nth' member of the target population.

4. Cluster Sampling, the selection of clusters (for example, classrooms, schools, districts) that contain individuals who comprise the object of study. The clusters should be selected as a probability sample, all of the individuals within a randomly selected cluster being included in the study.

5. Multi-stage Sampling, the use of various stages of sampling. For example, an education system is selected, then schools within the system are randomly sampled, classes within the schools are then randomly selected, and finally students within the classes are randomly included for study (Gay and Airasian, 2000; Krathwohl, 1997; Scott and Usher, 1999).

If a research document does not indicate that scientific probability sampling was used or actually states that this approach was not used, then the results of the research should be taken with reasonable scepticism. Without scientific probability sampling, the conclusions of the research, the generalizability of the results to the population of interest, and the potential for use in a research review outside the original setting, are all questionable.

There is a distinct and important difference between random sampling and haphazard sampling. Unfortunately, research will sometimes utilize techniques that are referred to as being random, but are in fact nothing of the sort. For example, a researcher might be positioned outside a secondary-school library to approach students as they exit the library and ask them a few questions. The claim might be made that since the researcher had no control over the specific student exiting the library at a given time, then the particular students interviewed must be the result of a random process. But the process is not really random. Many things influence when a student leaves the library (or even spends time in the library at all!), such as class schedule, appointments with teachers or administrators, or even employment schedule. None of these reasons for being in the library or leaving it at a certain time creates a condition of randomness. Any claim of randomization must be supported by the facts, not simply by assertion.

However, some educational settings, particularly those in rural areas of developing countries, lack the kind of records from which random procedures can operate. If the context of the research dictates, structurally, that randomness is not reasonably possible, then the results should not be disregarded. Flawed research, if structurally dictated, at least provides some information. However even unavoidably flawed research is still flawed, and it should be used with serious reservations and with care in drawing conclusions, particularly if it is to be applied regarding important decisions on policy and education.

## *Measurement*

The fifth element in the research review process is to evaluate the research measurement and the measurement instruments. In quantitative research, measurement indicates that some aspect of education, such as student learning and performance, impact of spending on learning, or some aspect of teaching, is being rendered in numerical form. The instruments used to measure may include formal standardized tests, surveys, interviews, checklists of materials and supplies, or regular government statistics regarding enrolment and attendance, among many possibilities.

However, almost all instruments used in educational research are assumed to be flawed in some way, either because the things being measured are difficult to render 'cleanly' in numerical form, or because the instruments used are in some way not an optimal 'fit' for the context of their use. Thus the reviewer must identify and take into account potential sources of measurement error: for example, poor instrument design, less than skilful administration of the instrument, inappropriateness of the instrument to the context, or any of a number of other challenges. Then the possibilities and potential meaning of errors in measurement must be considered in the research review.

Measurement is generally regarded as the process of quantifying the educational condition or variable of interest. For example, if the conditions of schools in a country are of interest, then the research would seek to quantify such things as the number and quality of books available, access to resources such as paper and pencils, or conditions of school facilities such as electricity, running water, and toilet facilities. If the performance of students is the condition of interest, the research might attempt to quantify how well students read, compute mathematics, or perform science experiments, among many possibilities. Measurement includes the use of instruments (for example, surveys, observations, tests, grades, entrance examinations, exit examinations), the compilation of data, and the attribution of meaning to the data.

Common types of measuring instruments assess such things as: (a) achievement, (b) aptitude, (c) attitude, (d) interest, and (e) personality. While traditional notions of measurement, and measurement error, reflect largely on the instrument being used in the research, there are other significant aspects of measurement that also impinge on quality. Since measurement is now taken as a process, it has become accepted to conceptualize measurement as everything that is done to arrive at some numerical value or estimate.

As the quality of measurement is being evaluated in each of the research documents being reviewed, the following questions need to be considered as potential sources of measurement error:

1. *How reliable and valid is the instrument being used?* Better research documents will describe the technical aspects of the instruments being used. Common indicators include instrument validity ratings, usually expressed in terms of correlation coefficients (*r* values), which indicate characteristics of the instrument such as content, criterion, and construct validity.

2. *How well trained were the researchers in administering the instrument?* Was each researcher extensively and systematically trained, and was that training, along with the administration of the instrument, a continuing element of the research project or a one-time event? The intra- and inter-rater reliability of the researchers involved in data collection would be important in asserting the quality of the results. Rater-reliability coefficients (*r* values) approaching 1.0 indicate much better training and reliability, and therefore less potential measurement error. Coefficients below 0.50, and certainly those approaching 0, would be considered unacceptable, and thus present a potentially serious source of measurement error.

3. *Was the sample or population being investigated homogeneous, or did it demonstrate wide variation on significant sociocultural characteristics?* An instrument's reliability and validity, and the justification for use of any instrument, can be seriously threatened as the complexity and variation in the subject group and the context increase. Sample and population complexity can stretch, or even exceed, the parameters of the original design of even the best of instruments and the expertise of the best of researchers in compensating and accounting for sample and population complexity.

4. *Were the operational definitions employed in the research precise or vague?* When operational definitions are vague, the 'fit' of the measurement to the problem can be seriously compromised; yet the problem is hard to detect because vague operational definitions render a broader range of outcomes as viable. This is a most subtle influence on measurement error, and one that takes careful thought to detect when evaluating research documents.

5. *Were the data used in a reasonable and scientific way?* All previous issues aside, if the ultimate use of the data was not within the limits of the instrument, its design and intended use, and within the normal considerations of scientific reasoning, then the results are highly subject to measurement error.

6. *Is there a standard error of measurement (SEm) indicated, and if so is it acceptable for the research being reviewed?* The *SEm* is an estimate of how often scores of a given size can be expected. A small *SEm* indicates high reliability, and a large *SEm* demonstrates low reliability. However, what represents a small or large *SEm* depends on the nature of the test. Since the *SEm* is expressed in the same units as the instrument, what represents 'small' is related to the unit of measurement of the instrument. For example, an *SEm* of 5 would be very large for an instrument with 10 items, but small for an instrument containing 250. Thus *SEm* must be interpreted carefully and within the parameters of the specific instrument for which it was calculated.

The instruments used for measurement, the process employed, and the ultimate use of the measurement data all impact the potential usefulness of the research results. All of the major aspects of measurement must be carefully considered before the results of a research document can be defensibly included in a research review.

## *Data analysis*

The sixth element in the process of research review is the evaluation of the data analysis. How data are treated in the analysis can influence the validity of conclusions as well as the usefulness of the study and results in other contexts. Sometimes simple omissions in the way the data were analyzed or in the nature of the data can lead to major problems.

For example, if the mean (or average) mathematics test scores of a group of students are presented without any indication of the distribution of those scores (for example, in terms of standard deviations) then misleading conclusions may be reached. Although the mean test score might appear very high and lead to positive conclusions regarding the mathematics competences of the students, the actual distribution of test scores could be bi-modal: two large clusters of students on either side of the average, one group performing very well and one performing quite poorly. If on average most students appear to be performing well, and the cluster of struggling students is not represented, obviously, serious flaws might occur in the policy formation process.

There is a significant difference between data analysis and data interpretation. Data analysis refers to the statistical treatment of the data, whereas data interpretation deals specifically with the meaning attributed to the results of the analysis.

The basic 'quality' of the data produced by research is determined largely by the rigour and quality of the general procedures, sampling, and measurement. If the reviewer has determined that these three areas are sufficiently well done, then it is possible that the ensuing data analysis will be of high quality as well.

The data analyses performed should be consistent with the stated purpose and problem(s) presented for the study. Thus if the study is descriptive, then the data analysis should be comprised of descriptive statistics (for example, mean, median, mode, standard deviation, variance). If the study is, according to its purpose and problem(s), a correlational study, then the data analysis should be comprised of correlational and/or regression procedures. And, finally, if the study is causal in design, then inferential statistics such as the *t*-test, ANOVA, and Chi-square tests should be used.

The reviewer must evaluate the fit of the data analysis procedures to the purpose and problem(s), as well as the overall design, for without internal consistency among these items the data analysis could represent no more than a search for statistical significance. Too often, researchers design a study to inquire concerning one topic and, finding no statistically significant results on the purpose and problem(s) stated for the study at the outset, engage in a massive 'hunt' for significance wherever it can be found. In these instances statistical analysis after statistical analysis are utilized in an attempt to find some, perhaps any, significant trend or pattern. When this situation occurs, the utility of the results of the data analysis is in serious question.

As the reviewer evaluates the data analysis section, each purpose and problem stated in the early part of the document should be matched to an ensuing analysis. The results of the analysis should be noted, and the appropriateness of the analytical technique to the corresponding purpose or problem should be assessed.

In assessing what can be concluded from the data analysis concerning a specific purpose or problem, the reviewer should keep an important issue in mind: statistical significance does not always equate to functional significance. At times the data analysis may identify a difference between variables or groups that is determined to be statistically significant, but when the size of the actual difference is reviewed it becomes clear that, in practice, the difference will not be functionally significant.

An example is found in intelligence testing. At times the general intelligence (IQ) of students is measured for various purposes, and often very large samples are used. A major issue in achieving statistical significance is the number of subjects included in the study. The larger the sample size, the more likely that small differences will be statistically significant. Frequently, studies of large samples will conclude that there is a statistically significant difference in IQ between groups of students, but a closer look reveals that the measured group differences are just a few IQ points, or in very large samples only a fraction of one IQ point. Often these differences are only a small fraction of a standard deviation or a small fraction of the *SEm*, at most. While the difference between groups, at this level, might be statistically significant, the functional significance in the classroom or over the course of a person's lifetime would be impossible to assert and defend. Hence 'significant' findings from the data analysis need to be assessed realistically.

When evaluating the data analysis, the reviewer should also consider whether sufficient information was provided to understand accurately the nature of the data. For example, descriptive studies may indicate the average score for a group (or some other measure of central tendency), but then present no indication of how the data cluster around that score (for example, standard deviation, variance). It is virtually worthless to be told that the average student score is, for example, 97 unless the distribution of scores around that average is also given. There is a huge difference between a group with a mean of 97 and a standard deviation of 2 (highly clustered around the average score) and another group with a mean of 97 and a standard deviation of 30 (loosely clustered, at best). These sorts of differences are more common than might be expected; therefore the reviewer should evaluate the information provided well beyond the simple declaration of statistical significance.

## *Results*

The seventh element in the research review process is evaluation of the results, particularly in the context provided by the project's stated purpose and problem(s). Clearly, all data analyses produce more results than are presented in the report. Sometimes the most potentially interesting or influential results are not presented in the report due to the filtering influences of the rationale and purpose of the original research.

However, the reviewer must carefully examine the results that are presented, particularly as they correspond to the stated purpose and problem(s). Often results are presented that are tangential, at best, to the declared purpose and problem(s). How the presented results 'fit' to the rest of the research project indicates, to a significant degree, the coherence of the project itself.

Not all research documents distinguish between data analysis and results or functions. When they do, the data analysis section tends to be somewhat perfunctory, presenting a fairly 'simple' and factual recitation of the analysis. In contrast, the results section, when separate, typically distinguishes between analyses that are significant in a technical sense and those that are relevant or central to the study itself.

Whether or not there is a separate section for results, the primary function at this point in the review is to match and reconcile the purpose and problem(s) with specific quantitative statistical procedures and results. All purposes and problems should be specifically and explicitly addressed, and the rationale for support or lack of support should be easily determined from the presentation.

When a purpose or problem presented early in a document is not specifically addressed in the data analysis or results section, a problem or limitation with the research is implied. And when a purpose or problem 'emerges' in the analysis without prior identification, then the research project has likely cast the statistical analysis 'net' wider than it should have done.

Evaluation of the results section is the most direct of all the evaluative functions. The main concern is matching purpose and problem(s) with the results of the analysis. Identification of any 'new' purpose or problem(s) beyond those originally stipulated in the document follows. Then an assessment of sufficient warrant of conclusions ensues.

## *Conclusions and generalizability*

The last element in the research review process is analysis of the conclusions and consideration of their potential generalizability to the specific policy context. As a combination of the preceding factors, careful consideration must ultimately be made concerning the reliability and validity of the conclusions made in the research articles, both individually and collectively. While individual articles might suffer from specific flaws, perhaps the collective strengths of these articles are reasonably substantial. Yet even though the collective strengths might indicate a solid basis for asserting a particular conclusion, that conclusion might not have reasonable applicability (generalizability) in the specific policy context under consideration. For example, a science initiative that has proven consistently successful in a number of studies in Western Europe and Japan might not be applicable to settings that have less resource accessibility, lack specific teacher expertise, or include greater population heterogeneity. Ultimately, the reviewer must determine, even when the research evidence is consistent, considerable, and coherent, whether those results have a reasonable application in the context being considered.

Although all of these eight elements will not be explicitly addressed for each research project included in the final review, the researcher should certainly consider and evaluate each element during the review process. The professional and moral obligations are clear: research reviews that influence policy formation must be rigorously conducted and rationally presented.

Review of the conclusions section of a document begins and ends with a simple question: are the conclusions warranted? The reviewer should evaluate whether each conclusion is, in fact, warranted by the data analysis. Specific and scientifically rigorous data analysis should be provided for each conclusion. Without sufficient warrant (evidence) a conclusion is not supported, no matter how much that conclusion confirms popular perception.

Generalizability has two components for the reviewer to consider: (a) generalizability from the sample to the original population of interest (internal validity), and (b) potential generalizability beyond the original population of interest to other populations (external validity). If the conclusions are the results of rigorous research design, as determined by an evaluation of the previous seven elements, then the generalizability to the population of interest is high.

But even high internal validity does not guarantee that the conclusions of the research will be usable in other settings. To determine generalizability beyond the intended population is difficult. A careful consideration of potential obstacles to the external validity of the research conclusions must be made. Beyond the match of the words and terms used in the purpose and problem(s) of the research to the setting of the research review itself, evaluation of the underlying issues must be made, guided by questions such as the following:

1. Is the nature of the educational condition under study in both the research being reviewed and that of the research review truly comparable?

2. Are the instruments being used appropriate, in a reasonably equal fashion, to both settings?

3. Are the differences between the populations of interest, as expressed in the characteristics of the research sample, sufficiently small for the data, results, and conclusions to be applied in both contexts?

4. Is the targeted context of the research review sufficiently similar to that of the research being reviewed to allow for reasonable transference of the results without the use of significant caveats or cautions?

If the reviewer finds numerous concerns, limitations and cautions concerning the applicability of the research to the setting of the review (for example, population, educational condition, and functional differences), then serious consideration should be given to the appropriateness of the research for the review. It is far better to be cautious about including research with limited potential for transfer of results than to write an overly optimistic review that could lead to unfortunate, unfounded policy in a critical area.

# V. Evaluating research

Quantitative research is designed to answer many different kinds of questions. In evaluating research, careful classification of the type of research being reviewed is critical in establishing both internal validity and external validity. While specialists in educational research may have differing classification systems, with some utilizing very fine-point distinctions among numerous typologies, most could be satisfied within the general system of three basic methodological categories: (a) descriptive, (b) correlational, and (c) causal.

Classifying research reports into the general categories of descriptive, correlational, and causal is one of the most effective initial procedures for evaluating research. Once this broad classification has been accomplished, the evaluation of each 'pool' of research within each category can begin. Typically, conclusions regarding the quality of the research within each area, and conclusions that are supportable from within that category, can be drawn. Then the research review will combine the results of these three 'area' reviews into one general presentation. The final research review is thus based on a careful analysis of each research document within the constraints of the type of research it represents, casting each one within the available pool of research projects that are similarly designed and therefore similarly constrained in terms of cause and effect potential.

## *Determining research methodology type*

The first task of evaluating each research document is to determine what type of research methodology is represented. The two primary approaches to this classification are: (a) to classify on the basis of the statistical procedures used, or (b) to classify according to the characteristics of research design. The easiest and most dependable classification process is to rely on the statistical procedures used in the report. However, some research will attempt conclusions not supported by the statistical analysis, having been designed to address problems beyond that analysis. This condition will be most typically found in research intended to be causal but containing design limitations. Therefore, even though classification can be typically based on statistical procedures, each research document should be reviewed in terms of research design as well.

### *Statistical procedures approach*

The three basic types of research methodology utilize specific types of statistical procedures and can therefore be categorized on the basis of these procedures. A reviewer does not need to be a statistician to engage in this activity, but the reviewer does need to be careful and thorough.

In using this approach the reviewer carefully examines the research document and notes the statistical approaches it presents. Fundamentally, there are two types of statistical procedures: descriptive and inferential. If a study presents *only* descriptive statistics, then it is highly likely, possibly certain, that the research is descriptive. When inferential statistics are presented, then the study is either correlational or causal.

Descriptive statistics present data oriented towards illustrating how the data 'cluster'. For example, if the research is reporting on the experience level of primary-school teachers, the statistics will likely present some indication of the average number of years a teacher has been in the profession. This is best done through the use of the arithmetic average, or mean. The mean, along with similar statistical measures such as the mode and median are called *measures of central tendency.*

In addition to measures of central tendency, a good descriptive research document will also indicate how the data are 'spread out' around the mean, median, or mode. Typically, this will be done with a statistical value such as a standard deviation, with the measure of central tendency being a mean. A standard deviation indicates how far away, on average, the scores are from the mean. In the case of the experience level of primary-school teachers, the average number of years in the profession might be a mean of 6.7 years, with a standard deviation of 2.1 years. From these statistics the reviewer would understand that the typical teacher has been in the profession for 6.7 years, but also understand through the standard deviation that most teachers have been in the profession for somewhere between 4.6 and 8.8 years (one standard deviation below and one above the mean).

Statistical indicators such as the standard deviation are referred to as measures of variability, or measures of dispersion. The most common measures of central tendency and variability are presented in *Table 5.1*.

**Table 5.1 Common measures of central tendency and variability**

| Measure | Purpose | Statistical symbol (if any) |
|---|---|---|
| Central tendency: | | |
| Mean | Arithmetic average of scores | For a sample = $x$<br>For a population = $\mu$ |
| Median | That score at which 50% of the other scores are 'above' and 50% are 'below' | |
| Mode | Score attained by more subjects than any other | |
| Variability: | | |
| Standard deviation | The average distance (deviation) of each of the scores from the mean | For a sample = $s$<br>For a population = $\sigma$ |
| Quartile deviation | Half of the difference between the upper quartile and the lower quartile in a distribution | |
| Variance | The spread among scores | For a sample = $s^2$<br>For a population = $\sigma^2$ |
| Range | The difference between the highest and lowest score | |

With descriptive statistics no correlational or causal conclusions can be made. Inferential statistics, on the other hand, indicate correlational and causal research and conclusions.

## Table 5.2 Common inferential statistics

| Statistic | Purpose | Statistical symbol (if any) |
|---|---|---|
| Correlational: | | |
| Pearson product correlation | Test whether a correlation/ relationship exists between two variables, using scale data | $r$ |
| Spearman's Rho | Same as Pearson's, using ordinal data | $\rho$ |
| Regression | Uses correlated variables to predict (not in a causal sense) the value of an outcome variable | |
| Causal: | | |
| Chi square test | Test whether variables are independent of each other | $\chi^2$ |
| *t* test | Test difference between means of two groups | $t$ |
| Analysis of variance (ANOVA) | Test difference between means of three or more groups | $F$ |
| Analysis of co-variance (ANCOVA) | Same as ANOVA, except adjusting for the influence of a control variable | $F$ |

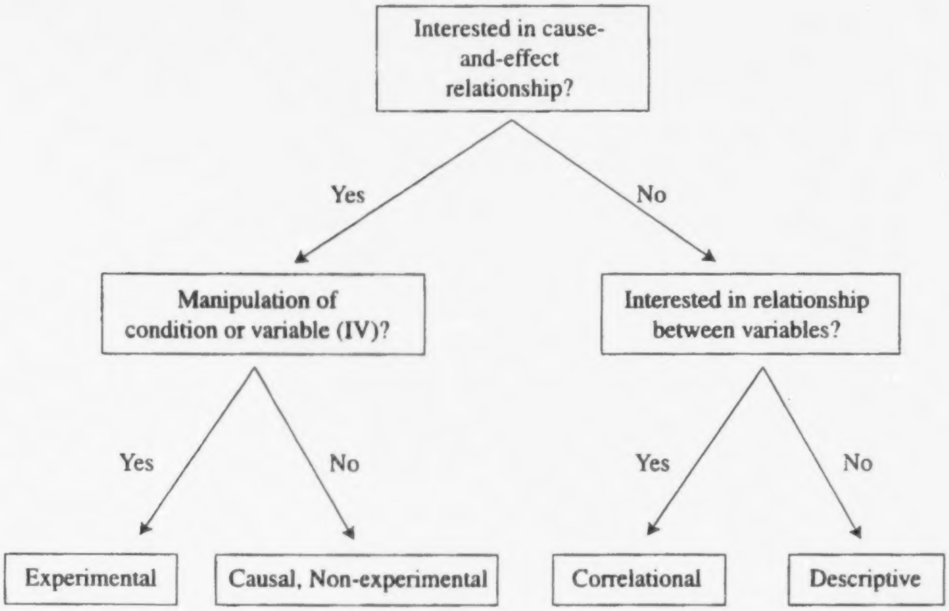
What can be confusing when first learning to determine the research type is that often correlational and causal research will also report and utilize descriptive statistics. The reviewer must keep in mind that it is the 'highest' level of statistics presented that identifies the type of research being conducted. The presence of 'lower' level statistics (for example, descriptive) does not make a research project simultaneously descriptive AND correlational, or descriptive AND causal. A causal research document is causal whether or not it also contains descriptive statistics. Typically, descriptive statistics are used as an introductory or foundational platform for the subsequent use of correlational or causal statistical procedures. The most commonly used inferential statistics are presented in *Table 5.2*.

In summary, the reviewer should evaluate the statistics reported in the research document, designating whether each one is descriptive, correlational, or causal. If only descriptive statistics are used, the research is most likely best classified as descriptive. If any correlational or causal statistics are used (regardless of whether descriptive statistics are also presented) then the research is correlational or causal. Once the type of research has been designated, then the research document can be evaluated according to the characteristics and limitations of each type of research.

### *Design characteristics approach*

The design characteristics approach is a more subtle way of establishing the type of research methodology that is being reviewed. With the single exception of establishing research as causal, there are usually only minor differences in design.

**Figure 5.1 Decision tree for identifying research type**


Interested in cause-and-effect relationship?
Yes
No
Manipulation of condition or variable (IV)?
Interested in relationship between variables?
Yes
No
Yes
No
Experimental
Causal, Non-experimental
Correlational
Descriptive

(Gay and Airasian, 2000, p. 22)

To identify causal research, the reviewer is looking for whether a direct manipulation of the condition under study took place. For example, if the research is designed to test the causal relationship between teaching style and student learning, then a causal research design would include the use of different teaching styles (the independent variable) and then measure the subsequent level of student learning (the dependent variable). If there is ever a direct manipulation of an educational condition (an independent variable), then the research is typically designed to measure the causal impact of that manipulation on an outcome (dependent variable).

Research that does not indicate the manipulation of an independent variable is most likely descriptive or correlational. Here the design characteristics become more central to the designation of research type. A reviewer must examine items in the report such as the purpose, problems, questions, and hypotheses. Not all research documents will include all of these elements, but those that are present need to be examined. *Table 5.3* presents examples of these elements in descriptive and correlational form.

**Table 5.3 Examples of purposes, problems, questions, and hypotheses**

| Descriptive purpose | Correlational purpose |
|---|---|
| To explore the current condition of the education system in Zimbabwe. | To explore the relationship of educational conditions to educational outcomes in Zimbabwe. |
| There is insufficient data on the conditions of the schools in Nepal, relative to the numbers of books and other basic educational supplies available to children in the most rural settings. | The relationship between basic educational supplies and the academic achievement of children in the most rural settings in Nepal must be better understood before consideration of a new national policy on resource allocation can be rationally made. |
| What percentage of primary-aged female children in Pakistan are enrolled as students in government schools? | What is the relationship between student gender and primary-school enrolments in Pakistan? |

| | |
|---|---|
| It is hypothesized that there is an positive increase in the number of six to twelve-year-old girls enrolled in primary education in Morocco from 1980 to 2000. | It is hypothesized that there is a relationship between the number of female teachers and the increase in enrolment of six to twelve-year-old girls enrolled in primary education in Morocco from 1980 to 2000. |

Often a research document will declare the type of research intended; however, this indication needs to be approached carefully, as many individuals engaged in research are not necessarily experts on methodology and design. There is always a strong possibility that the author(s) of a research document have used a term like *descriptive*, *correlational*, or *causal* in a vague, or perhaps inaccurate, way. Declarations of research type by the author(s) can be used as clues to the type of research being presented, but concurring evidence must always be sought.

**Table 5.4 Examples of correlational and causal language**

| Correlational language | Causal language |
|---|---|
| Student gender is ***related to*** performance on mathematics tests. | Scores on mathematics tests are ***caused by*** teacher experience, style and knowledge. |
| Success in college and university studies is ***predicted by*** level of performance in secondary school academic subjects. | Time spent on non-academic pursuits, such as work and social activity, ***leads to*** lower academic performance in university and college studies. |
| Financial and socio-economic status of student families is ***associated with*** the quality of educational opportunity of students. | Access to financial and material resources ***determines*** the quality of educational opportunity provided for students. |
| The amount of time teachers interact with students individually is ***correlated to*** the number of students in the class. | Increasing the number of students in a class ***results in*** a reduction of teacher interaction with individual students. |

Finally, the conclusions presented at the end of the research document will give strong indications of what type of research methodology was involved. The words used in presenting conclusions or making assertions from the data demonstrate the intentions of the research. The most direct conclusions can be drawn from correlational and causal research. *Table 5.4* presents examples of common correlational and causal language. The salient words indicating correlational and causal language are presented in italics.

### *Descriptive research*

The most common type of educational research is descriptive. When the purposes and problems of a research project are oriented towards simple exploration or clarification of the current status of an education system or unit, descriptive research is the design of choice.

Typically, descriptive research is a first step leading to more complex questions. Examples of descriptive research are readily found in the efforts of national educational ministries to track student enrolment information, gather information regarding the condition of schools in terms of numbers of textbooks and other basic school supplies, or describe the levels of teacher education and experience, etc.

Descriptive research is the least intrusive form of research and is the most easily conducted, in terms of political support and necessary research expertise. However, it is limited in the complexity of purposes and problems that can be supported scientifically. For example, while descriptive research can reveal how many students are enrolled in schools in Finland, it cannot explain *why* that many students are enrolled or why there are more or fewer students enrolled this year than the last.

This is not to say that descriptive research is not useful. Most of the types of research designed to explore more sophisticated purposes and problems depend on solid descriptive research as a prior step. However, research documents utilizing more sophisticated methodologies must be found if purposes and questions beyond basic description are sought.

Not only does descriptive research explore how things are, descriptive research may also seek to determine how groups of people (for example, men and women, teachers and parents, government school students and non-government school students) view issues and topics. Descriptive research represents a high proportion of the available research, and thus is likely to be the most common type of research reviewed. Therefore, the characteristics and limitations of this type of research are critical to understand.

The two most common forms of descriptive research are those produced by standard educational 'audits' (for example, recurring national, regional, or system-based accounting of student enrolments, conditions of facilities, and access to materials and supplies) and by surveys, questionnaires, and interviews. Although surveys, questionnaires, and interviews can be used in conducting audits, they are regularly utilized for much more.

Typical descriptive educational research studies measure the conditions of education through audits, and assess attitudes, opinions, preferences, demographics, practices and procedures in education through the use of surveys, questionnaires and interviews (Gay and Airasian, 2000, p. 275).

A common characteristic of descriptive studies is the extensive use of tabular and graphic displays. The following are the most commonly used methods of dispiay:

1. *Frequency tables* present descriptive statistics (mean, median, mode, standard deviation, variance, range, percentage/proportion distribution, etc.) in simple numeric form.

2. *Graphic tables and figures* present descriptive data in various graphic formats that demonstrate relationships – comparisons and distributions – in a set of data in a variety of formats:

    (a) *Line graphs* in which the lines connecting data points show the values of one or more variables over time or conditions, with two or more variables plotted so that comparisons between trends can be easily made;

(b) *Bar graphs*, which are used when the independent variable is categorical data, with the height of each bar showing the value or frequency of the data;

(c) *Pie graphs*, which show the percentage or proportion, with the angle of the pie 'slice' from the centre being proportional to the frequency;

(d) *Histograms*, which express continuous data in bar form, representing relative frequency in the data set;

(e) *Frequency polygons*, which join together the mid-points of the bars of a histogram;

(f) *Scatter plots*, which express on a two-dimensional graph the values of single events/individuals plotted on two variables, with meaningful clusters of plotted dots implying correlations (American Psychological Association, 1994, pp. 142-163; Scott and Usher, 1999, p. 71).

Descriptive research is intensive in the presentation of data but appropriately cautious in the use of the data to draw conclusions beyond simple declarations of fact; therein are its advantages and limitations.

*Advantages and limitations of descriptive research.* Since descriptive research typically uses data that are regularly collected by educational institutions and governments (audits) or extracted from surveys, questionnaires, and interviews, it is generally an efficient and relatively inexpensive way of obtaining large quantities of data. Herein lies the main advantage of descriptive research: ready access to plentiful, relatively inexpensive data, allowing the descriptive presentation of information on an extensive range of educational issues and conditions.

Thus descriptive research provides information on a wide range of educational issues and conditions to a broad spectrum of consumers. Considering that most non-researchers are exposed daily, through newspapers and television broadcasts, to the various forms of tabular and graphic representations utilized in descriptive research, it is the most accessible form of research to interested consumers ranging from educational research experts to ordinary citizens.

However, the promise of descriptive research is laced with limitations. The most significant of these is the fact that simply describing a particular trend in enrolments, opinions on the quality of education, or conditions in schools does not mean that any conclusions regarding the cause of those trends, opinions, questions, or relationships among them are supported by the data. Yet many consumers and presenters of descriptive research attempt to do just that – assert either a correlational or causal link in descriptive data.

The limitation that no scientifically defensible correlational or causal assertions are possible from descriptive research is substantial. As earlier stated, educators, funders, parents, students and government officials all want to make good decisions informed by defensible scientific evidence. They prefer to make those decisions on the basis of scientifically proven cause-and-effect relationships, or at least within the realm of scientifically defensible correlation research. However, the most common and easily obtained research data derive from descriptive studies, and these cannot support correlational or causal inferences.

When a reviewer has identified a research document as descriptive, this limitation must be of the utmost concern. While sound description of educational conditions may be contained in the descriptive research being reviewed, neither the research document nor the reviewer is justified in making any correlational or causal conclusions. When the author(s) of a research document begin(s) asserting inaccurate conclusions, the reviewer is responsible for noting these instances, considering the impact on the usability of the research in the review, and accounting fully for inadequately grounded assertions in the final review.

However, none of these limitations, or potential misuses of descriptive research, should be considered a fatal limitation on the potential utility of these research documents. In fact, descriptive research is very valuable in identifying potential areas for further correlational or causal investigation. Its value is also undisputed as

BEST COPY AVAILABLE

support for correlational or causal research being reviewed. The reviewer simply must make certain that the limitations of descriptive research are not ignored in constructing a reasonable and defensible review of research literature.

### *Correlational research*

The first 'leap' in research methodology is seen in designs intended to answer questions of correlation between conditions (variables) in educational settings. When a research project undertakes to establish that there are relationships (which is another way of saying correlations) between two or more conditions in education, studies are designed that can explore and establish these relationships statistically.

Many of the same types of data used in descriptive research are found in correlational studies. However, the analysis of the data is accomplished in such a way that relationships between variables become both more evident and more scientifically defensible.

For example, while a descriptive study might simply present enrolment information and family income data for a group of students in Bolivia, a correlational design will statistically explore ways in which variations in enrolment 'match' the variations in family income in systematic and predictable ways.

Even though correlational research allows the defensible assertion of relationships between variables, it does not allow the identification of causal links between those same variables. To clarify causal links in a scientifically defensible manner, research must take an additional step in design and statistical analysis to become causal forms of research.

Correlational research can be considered descriptive, because it does describe a current condition. However, in describing conditions, correlational research expressly attempts to explore and establish whether relationships between two or more quantifiable variables exist and, if so, to what degree they exist.

If correlational research establishes a relationship, it means that scores or values on one variable are demonstrably associated with

BEST COPY AVAILABLE

scores and values on another variable. For example, it has been demonstrated over time that there is a relationship between academic performance in secondary school and academic performance in university. Secondary students who perform well in their academic subjects tend to go on to perform well at the tertiary level, and students who struggle in secondary school tend to struggle academically in their university studies. This type of relationship is considered to be positively correlated: that is, the association between the two variables of secondary school academic performance and tertiary-level performance is typically in the same direction - when one is high, the other is high, and when one is low, so is the other.

Another approach in correlational research is to take variables that have been established as related (either positively or negatively) and attempt to use them to predict another outcome variable. In the example of secondary-school performance and tertiary-level academic performance, one might build upon the consistently established relationship and use one (secondary performance) to predict the other (tertiary performance). Other variables, such as scores from university entrance examinations or secondary leaving examinations, might be added to secondary-school performance to enhance the predictive power on tertiary performance. These predictive correlational statistical procedures are called regression analyses.

In both correlation and regression a number of potential variables are typically used, and only those that demonstrate higher levels of significant relationship are retained. Determining what numerical values qualify as 'high', however, can be challenging.

As the reviewer evaluates correlational and regression research, the reported coefficients need to be reviewed and assessed in terms of usefulness and meaning. This is the most challenging aspect of evaluating correlational research. What may be considered significant by one researcher may appear weak to another. Since most correlational techniques render the statistical values as scores between 0 and 1, a few examples are given in order to demonstrate the challenges in interpretation.

For example, correlation research may report that the relationship between student academic achievement and level of parent education is strong, with $r = 0.41$ being the correlation coefficient upon which the assertion was made. What this means, in general terms, is that approximately 17 per cent of the variance (calculated as the square of the correlation multiplied by 100) observed in student academic achievement is accounted for by variance in the level of the student's parents' education. While it may not seem like 17 per cent is a very high figure, it is more impressive when considering that there are numerous influences on how a student performs academically. When one variable among dozens is found to account for 17 per cent of the variance, the number becomes more impressive.

As a contrasting example: correlation research might report that there is a significant relationship between the number of years a teacher has spent in a particular school and the number of years the teacher has been in the teaching profession. At first consideration, a correlation of 0.9 might seem impressive and highly significant. However, this is an example of a structural relationship. Of course teachers who have been teaching for many years have been in a particular school longer than newer teachers; they have had more years of teaching to be in any school, let alone a specific one. It would be simply impossible for a person who had taught for only one year to have been at a particular school longer than one year, and a teacher who has taught for 30 years may very well have been in the same school for all of those 30 years. Therefore, in this example even a correlation coefficient of 0.90 is not particularly meaningful.

These two examples demonstrate that careful consideration must be devoted to the context of the reported correlation coefficient, and that the actual evaluation and interpretation of any numerical value must take into account the range of potential contextual influences. It is an interesting fact that a correlation of 0.41 can be considerably more meaningful than a much higher correlation of 0.90, all depending on the context and issues associated with the variables being correlated.

*Advantages and limitations of correlational research.* Correlational research is particularly popular in education because it often uses data that are already collected or that are readily available: for example, student performance scores, school resources, social and personal demographic information, and exit examination scores. Given that so much information is readily available and therefore costs very little to obtain, correlation research is popular among researchers and educational institutions interested in more than simple description. Correlational research is also typically less time-consuming to perform than causal research designs, which usually require longer-term, longitudinal intervention.

Additionally, correlational research can investigate potential relationships between many variables, typically with no particular risk to the students involved. This is important because experimental research, which is more difficult and expensive (both financially and politically), can be guided by the results of correlational research. In fact, the role of correlational research in guiding potentially fruitful experimental research is one of the primary benefits of correlational work.

The relationship of efficiency between correlational work and subsequent experimental research is easily clarified. If two variables are correlated, then the possibility exists that those same variables *might be* causally linked. But if two variables are shown through correlational research to be unrelated, then they *cannot be* causally linked. Thus much more expensive and time-intensive experimental research that might have been based on non-correlated variables that cannot be causally linked can be avoided, and limited research resources can be directed to the investigation of possible causal relationships between variables that have been established, through quality correlational work, as potentially causally linked.

Since this potential for causality is indicated for variables shown to be correlated, there is an almost irresistible inclination to establish, or at least strongly imply, causal links. This is the greatest risk in correlation research, especially if causal research, for any of a number of reasons, might not be possible.

However, the reviewer must remember that correlation only indicates the potential for causality; it in no scientifically defensible way establishes causality. Even if the author of the research indicates causality, it is the professional responsibility of the reviewer to carefully redirect the conclusions drawn from correlational research to the much more tentative condition of indicating relationships, correlations, and prediction.

A distinction needs to be made between prediction and causality. In the example of secondary-school performance as a predictor of tertiary-level performance, it is typically accurate to say that high academic performance indicators (for example, grades, test scores) in secondary school can be expected to predict commensurately high academic performance indicators (for example, grades, test scores) in tertiary education. However, it would not be defensible, or even rational, to assert that secondary performance indicators *cause* tertiary performance indicators. Test scores do not *cause* other test scores. A consistent pattern of relationship between test scores does not establish that those scores are causally linked.

### *Causal research*

All research aspires, in one way or another, to identify causal links between educational efforts and subsequent educational outcomes. If policy and funding could be based on clear, precise and defensible knowledge of exactly what efforts or expenditures would invariably lead to specific results, decisions in education would be much easier to make, and the results would conform more consistently to the intended purpose.

Causal research in education is typically one of two kinds: (a) that research which purposefully controls and manipulates the conditions of education to explore subsequent results, or (b) that research which employs cause-oriented statistical analyses of educational conditions that were not, could not be, or should not have been manipulated. This will be discussed in greater detail later.

An example of causal research would be an exploration of the impact of a particular reading curriculum on subsequent reading ability. If one group of students in Siberia is taught according to a 'new' reading curriculum and another group receives instruction under the 'typical' reading curriculum of the school, then the condition of reading instruction has been manipulated. If there is a difference between the subsequent reading ability of the two groups of students, then some level of causal conclusion would be supported by the research. Of course, there are many design and analysis issues involved in determining how scientifically defensible this causal conclusion would be, but the process illustrates how educational conditions are manipulated for a causal link to be established.

However, the complexity of the real world of schools and classrooms usually makes the results of even causal research somewhat unreliable. There are simply more environmental influences in educational institutions than would be found in a highly controlled laboratory setting. Nevertheless, causal research, if available, facilitates a more assertive stance in development and evaluation of educational policy than any other form of research.

If researchers had the option, and if all costs (for example, financial, political, social, moral, methodological) were equal, then reliable and valid causal research would nearly always be preferred in a review. A high-quality causal research document is the most potentially valuable source of information because not only are the educational conditions of interest described and the correlational nature of the variables demonstrated, but the actual nature of their causal relationship is known, within reasonable constraints.

However, causal research is often difficult to locate, and high-quality causal research is rare. This rarity is often related to the fact that educational settings and issues are complex, and many are laced with challenges that may preclude the use of a number of even basic, essential, causal research design elements.

For example, one of the most fundamental assumptions of most high-quality causal (experimental) research is true randomization. Technically, randomization assumes both random *sampling* from the population and random *assignment* to the condition(s) within the experimental design. In many educational settings, if not most, true randomization in sampling can be difficult to implement.

For example, in the USA the majority of students attend schools based on where they live. To set this system aside and randomly choose students from the population regardless of where they live, and then assign them to particular schools where they would receive the 'treatment' being researched, would be virtually impossible from a political and legal standpoint. Thus in the USA the educational treatment under study is often assigned randomly within pre-existing school and/or classroom populations. This form of 'randomization' presents serious challenges to the reliability and validity of the research project and results. Many causal research documents are limited on the basis of challenges, limitations, and even flaws in design. However, limitations aside, causal research of varying degrees of quality is frequently encountered in conducting a review of research literature.

*Advantages and limitations of causal research.* In reviewing causal research, the issues of internal and external validity become more salient than in other forms. Because causal research aspires to produce more declarative and authoritative conclusions and to establish causal links between independent variables (IVs) and dependent variables (DVs), the need to evaluate the quality of the research itself (internal validity) and establish its potential usefulness outside the original setting (external validity) is heightened. When policies are made on the basis of causal research and significant funds are committed, scrutiny must be intensified. While the potential for benefits is higher, so is the possibility of serious long-term mistakes.

There are basically two types of causal research: (a) research conducted after the actual 'exposure' to a treatment or condition, and (b) research that actually identifies and manipulates the treatment or condition with the sample. Regardless of the type of causal research, there will always be at least one independent variable (IV) and one dependent variable (DV) under study. An IV is the element or educational condition that is believed to be the cause of a specific outcome. The outcome of the exposure of a sample to the IV is called the dependent variable, because its value is determined by, is caused by, or is *dependent* on the impact of the IV. The assertion of cause and effect is made both by presenting the research design elements that support a causal claim and by specifying the strength of the asserted causal link through statistical analyses, such as $t$-tests and ANOVAs.

In the first type of causal research, where the treatment or exposure to a treatment or condition has already taken place before the research is undertaken, statistical analysis provides the primary basis for supporting the claim of causality. This type of causal research is often called causal-comparative research, or *post hoc* research (Gay and Airasian, 2000; Krathwohl, 1997).

In *post hoc* causal research there has been no manipulation of the IV in the research. The reasons for the lack of manipulation are multiple, but typically the lack of manipulation results because either the IV could not be manipulated or should not be manipulated.

The two most common situations where IVs cannot be manipulated are: (a) when the change in educational condition (IV) happened before the researcher was aware and was able to take pre-manipulation measurements, and (b) when the measurement of the manipulation should not take place due to the nature of the IV being manipulated, or due to legal and ethical constraints imposed on purposefully manipulating the IV.

Often educational curriculum, conditions in schools, etc. take place without the involvement of research personnel who could establish the conditions necessary to conduct a rigorous experimental design. Schools will often purchase new textbooks, introduce a new way of teaching, or improve the quality of teaching staff, etc. and then notice a change in student outcomes that appears to be significant. In these cases researchers must attempt to assert that the new textbooks, for example, are the cause of the observed improvement in student performance.

Establishing scientifically defensible assertions of cause and effect, in these sorts of circumstances is difficult. Often research will attempt to replicate an experimental design that actually controls for the introduction of something such as textbooks to one group of students, but not to a similar group of students, by finding a 'matching' group of students who did not receive the same new books. By statistically comparing the scores of the students who had new books against those who used the other books, a research document will attempt to assert causality.

Clearly, this procedure and others that attempt to approximate direct control and manipulation of the IV are inferior to true experimental procedures and designs. While the same statistics are used in these studies that are used in experimental studies, the results are not as reliable.

When reviewing such *post hoc* studies, the reviewer must take careful account of how the lack of direct control and manipulation might threaten the validity of the findings. Procedures designed to approximate randomization, direct control and manipulation, pre- and post-testing of the subjects, etc. are all less desirable substitutes for real experimental designs, and the results of these studies should be used with great caution.

The greatest concern in *post hoc* causal studies is the difficulty in asserting ***antecedence***. Antecedence is perhaps the most important element in establishing a defensible causal link between an IV and DV. This term refers to the fact that the researcher must be able to demonstrate with reasonable confidence that the DV actually followed the introduction of the IV. For example, if the change in performance of the students cannot be easily shown to follow the introduction of the textbooks, then it cannot be defensibly demonstrated that it was in fact the textbooks that are responsible for the *antecedent*, or subsequent improvement in student performance. With research after the fact, there are often many substantial challenges to the clear assertion of antecedence, since the change in DV could have begun before the introduction of the IV and as a result of some other unidentified IV.

The second type of causal studies, experimental designs, are much more scientifically sound and reliable. These are studies that identified the IV and DV of interest *before* manipulation occurred. As a result, experimental designs are typically much more dependable in internal validity and in the potential for external validity.

Often experimental designs employ some type of random sampling, a measurement of the educational condition to be manipulated (IV), controlled manipulation of the IV, a measurement after the manipulation, and the selection of a control group that is comparable to the experimental group but was not exposed to the manipulation of the IV. The results of this kind of study are possibly the best type of information to be found for policy formation.

However, the more 'control' utilized in a study (for example, randomization of the sample, the use of a laboratory school rather than an actual school), the less like the 'real world' an experimental design becomes. Increased control over the experimental condition increases the internal validity of the study, but decreases the external validity at the same time. Thus more highly controlled experimental studies will produce more scientifically defensible results for the sample being examined, but will also produce results that are far more difficult to transfer to educational conditions and systems outside the original target population.

## *'Gaps' in educational research*

Although the research reviewer will conduct an exhaustive search for relevant materials, it must be kept in mind that almost all educational topics have 'gaps' in the research literature. These gaps most often cluster around the type of research that has been conducted. *Table 5.5* presents theoretical, somewhat exaggerated, examples of what might be found in a thorough review of the research literature.

**Table 5.5 Examples of typical 'gaps' in the research literature**

| Research | Topic of interest | | | | |
|---|---|---|---|---|---|
| type | School size and academics | Family income and achievement | Gender and access | Literacy and productivity | Spending and efficiency |
| Descriptive | None | 13 studies | 51 studies | 33 studies | None |
| Correlational | 50 studies | 27 studies | 97 studies | 64 studies | 119 studies |
| Causal | 3 studies | None | None | None | 15 studies |

BEST COPY AVAILABLE

As just described, different types of research are designed to provide evidence to answer specific kinds of questions. Some research is better suited for questions of cause and effect, while other approaches to research can only answer questions concerning non-causal relationships.

During the search for research documents, the reviewer must be aware of these gaps and be prepared to account for the limitations in the types of conclusions that can be reached if those gaps appear in areas impacting the answers that are being sought in the policy process.

Gaps in available research indicate where future work is most needed. Gaps also provide clues to structural limitations that might not be open to solution. For example, high-quality experimental research, which is the most reliable and valid way to establish cause and effect, requires the manipulation of the possible causes for the effects of interest. A number of substantial 'causes' in education, such as teenage pregnancy, family income, racial and ethnic group membership, etc., are not open to direct manipulation. Therefore, high-quality experimental research may not be available for certain important questions.

Gaps do not necessarily indicate that important information cannot be gathered regarding the topic of interest or that the collective body of research on a topic is somehow 'flawed'. When gaps occur, the reviewer must identify them and examine their meaning and implications. Good research reviews always indicate where these gaps exist and make use of them in substantial ways.

# VI. Synthesizing and presenting research analysis results

Following analysis of the technical merits of the individual research documents, the work of crafting a meaningful synthesis of the collective body of research must be undertaken. A synthesis goes beyond the presentation of the 'facts' of the research documents to make sense of the merits and limitations of the body of research being evaluated. This component of research review is the most challenging because it represents much more than recitation or summary. A quality synthesis requires thoughtfulness, professionalism, and political savvy.

## *Synthesis versus serial recitation*

The reviewer should always keep in mind that what is being created is a *review* of the research literature, not simply a serial recitation of what others have written. Not all authors and/or research will be in total or even partial agreement. Since the body of research being reviewed will be complex, the author of the research review must try to make sense out of the conflicts or potentially confusing inconsistencies in the literature. This synthesis is the primary contribution of the author in the research review.

The most difficult aspect of reviewing research literature, particularly for a novice or beginner, is to actually assert not only what was found in the research literature reviewed, but to interpret what it all means. Traditionally, research was presented in a 'dispassionate' voice, removed from the real world of application. However, in more recent years it has become not only acceptable but expected that the results of research will be couched within the real world of educational application, political and policy complexity, and financial challenges and constraints. Therefore, the reviewer must attempt to create a new perspective supported by the positive and negative evidence presented in the various research documents.

A synthesis is thus a new creation crafted from pre-existing materials. And just as research documents were examined for a purpose and problem(s), the research review must be oriented and guided by its own purpose and problem(s). These should be presented very early in the review, and the review will focus on addressing them directly. In fact, addressing the purpose and problem(s) of the review of research should form the superstructure of the review.

In creating the research review it is best to first develop a topic outline that presents the 'argument', or line of argument, that the review will take. This can be presented as an assertion or a proposition or as a group of assertions and/or propositions at the beginning of the review. Then the review presents an extended narrative that demonstrates and warrants the argument with direct quotations, citations, and graphic presentations.

## *Narrative presentation*

The narrative is the body of the review of research. It is best to begin writing the narrative by constructing an outline of the presentation, complete with a list of intended quotations, citations, and graphic presentations. A sample outline, which is adapted from the working outline used in producing a published book chapter (Atchoarena and Hite, 2001), is shown in *Figure 6.1*. It represents the first stage of outlining for a narrative presentation.

The writing of the narrative could actually be viewed as 'filling in' the outline through repeated revisions and editing. Of course, the structure of the final document may deviate from the initial outline, but using an outline as the working model of the review provides coherence to the undertaking, by forcing the reviewer to consider how particular components of the review 'fit' together in a structured way.

Within each level of the outline for which there is a reasonable body of research, the reviewer should plan to present the results and conclusions drawn both from individual studies and from the entire set of studies that are relevant to the specific topic. The conclusions drawn across studies should represent the major portion of the narrative, while issues, examples, and specifics are illuminated by periodic references to individual research documents.

**Figure 6.1 Sample research review outline on educating economically disadvantaged people in Africa**

---

- 1. Economically disadvantaged people in Africa: the scope of the problem.
  - 1.1. Describe context in ways that are very general and common to all countries.
    - 1.1.1. Economic (including labour market, urban *v*. rural, etc.).
    - 1.1.2. Demographic.
    - 1.1.3. Social (discuss the issues of poorly educated *v*. disadvantaged).
  - 1.2. Describe the political background.
    - 1.2.1. Rise of social partners.
    - 1.2.2. Rise of labour organizations.
  - 1.3. Brief educational system summary.
    - 1.3.1. Present system, discuss issues of illiteracy and enrolment rates.
    - 1.3.2. Investment in education.

- 2. Review reforms in terms of basic education, policy trends.
  - 2.1. Reforms of school systems in sub-Saharan Africa.
    - 2.1.1. Structural reform.
    - 2.1.2. Curriculum reform.
    - 2.1.3. Technological reform.
    - 2.1.4. Community participation reform.
    - 2.1.5. Decentralization reform.
  - 2.2. Non-formal education.
    - 2.2.1. Alternative strategies to basic education: labour market integration.
    - 2.2.2. Modernizing traditional apprenticeship systems.
    - 2.2.3. Adult education and local development.
    - 2.2.4. Trends in the provision of literacy and vocational skills to adults.

- 3. Main achievements and obstacles.
  - 3.1. Progress and resistance in the school system.
    - 3.1.1. Access – has it increased?
    - 3.1.2. Participation – has it increased?
    - 3.1.3. Efficiency – has it increased?
    - 3.1.4. Drop-outs and their impact on the labour market.

- 3.1.5. Equity, especially for girls, and has the gender-gap decreased?
- 3.1.6. Quantity v. quality.
- 3.1.7. Conditions of teaching and learning.
- 3.1.8. Measures of achievement/attainment, are outcomes improving?

3.2. Hopes and questions regarding non-formal education.

- 3.2.1. Do non-formal education programmes provide a meaningful and sustainable alternative or complement to formal education?
- 3.2.2. Is there an effective relationship between education and skills training?
- 3.2.3. Is there an effective partnership among non-formal education providers and with Ministries of Education?

4. Conclusion: Policy challenges to increasing educational opportunities for economically disadvantaged people in Africa.

In writing the narrative the reviewer must be careful to avoid presenting 'just the facts'. This approach to reviewing research is not only outdated, but it ignores the obvious reality that facts do not 'speak' by themselves. The reviewer chooses which facts are presented, in what order, and with what emphasis. It is fully the responsibility of the reviewer to narrate why certain facts are being presented and to clarify the meaning and role of those facts in the overall narrative presentation.

The narrative presentation is written 'to' an audience chosen by the reviewer. The task of synthesizing and presenting the results of the review must be focused on a specific audience. This focus is critical because it influences the type of information presented and the format of the presentation. Analyzing the prospective audience allows the reviewer to adapt to the level of education, sophistication in educational matters, primary policy interests, professional orientation, etc. of the group.

The reviewer also needs to consider the 'person' in which the review is written. While it has become popular in some literature forms to write in the first person, research literature reviews are still expected to be written in the traditional third person. Almost all publication manuals provide extensive examples of how this should be accomplished. If there is any question on 'person', the reviewer should consult any of the standard manuals (American Psychological Association, 1994).

### *Essential v. important v. interesting information*

The results of a properly conducted and thorough literature search will produce, hopefully, a significant number of research documents for the review. In the age of computers and inter-library loans systems, the amount of information can often be overwhelming.

In dealing with what seems to be an intimidating amount of information, the reviewer must adopt a framework for sorting out what is necessary to include in the review and what can be reasonably omitted. One such mechanism is to classify each document, conclusion, result, theme, or other component as: (a) essential, (b) important, or (c) interesting.

*Essential* information is what a member of the target audience for the review *must* know or be aware of in order to avoid being at risk. There are certain fundamental facts without which an involved consumer or participant would be unable to make properly informed and defensible decisions. For example, if the subject of the review is increasing access to basic education in developing countries, a consumer would be dangerously uninformed if the issues surrounding gender-based barriers to equal access were never introduced. The challenges facing young girls in gaining access to basic education are so fundamental to this issue that every review on this subject should address these challenges directly, and a consumer would be seriously under-informed by its omission.

Typically, essential information is first presented in the introduction, as some readers of research literature reviews may read only the introduction and conclusion sections of a review. Therefore, the reviewer would place the essential information in the section most likely to be read. Another way to conceptualize essential information is to use themes of these essentials as the focus of the body of the review. If in the introduction the essentials are presented in shortened, concise form, the body of the review is a place for each of the essential themes or points to be given more explicit discussion. The introduction thus sets up the 'skeletal' structure for the body of the review.

*Important* information includes what the consumer of the review *should* know in order to be well informed on the subject. Important information provides the supporting evidence and elaboration for the essential information presented in the introduction and developed in the body of the review. The important information is the basic substance of the argument or presentation, and without it the essential points would remain as relatively unfounded assertions, rather than supported findings and conclusions.

*Interesting* information typically includes what is of interest to a specific researcher or the reviewer, but is not needed by the consumer in order to make a well-founded conclusion. Interesting information could be considered secondary commentary, and thus is neither essential nor important. This type of information should be a minimal part of the product, if not excluded altogether. Unfortunately, many reviews consist mainly of what seem to be endless presentations of information that is not central to the purpose and problems that should form the review. A good reviewer will reduce significantly information that is interesting but not required to accomplish the task at hand.

The best reviews of research literature are concise and efficient, and thereby effective. Most consumers of research reviews are not content specialists or experts, and they therefore have little interest in minute detail and seemingly endless background information. Review and present the essential and important information, and reserve the interesting information for a different kind of endeavour.

### *Building an argument: the inverted triangle*

One way to consider the presentation of the results of the research review is to conceptualize an inverted triangle. If the structure of the presentation begins at the base of the triangle and proceeds towards the tip, a narrowing, or focusing, takes place. This is largely how a research review document would proceed in its presentation. The introduction provides a very broad overview of the review and then proceeds to work from the broadest issues and elements down through increasingly focused topics until the final, very focused conclusions are presented.

**Figure 6.2 The inverted triangle**

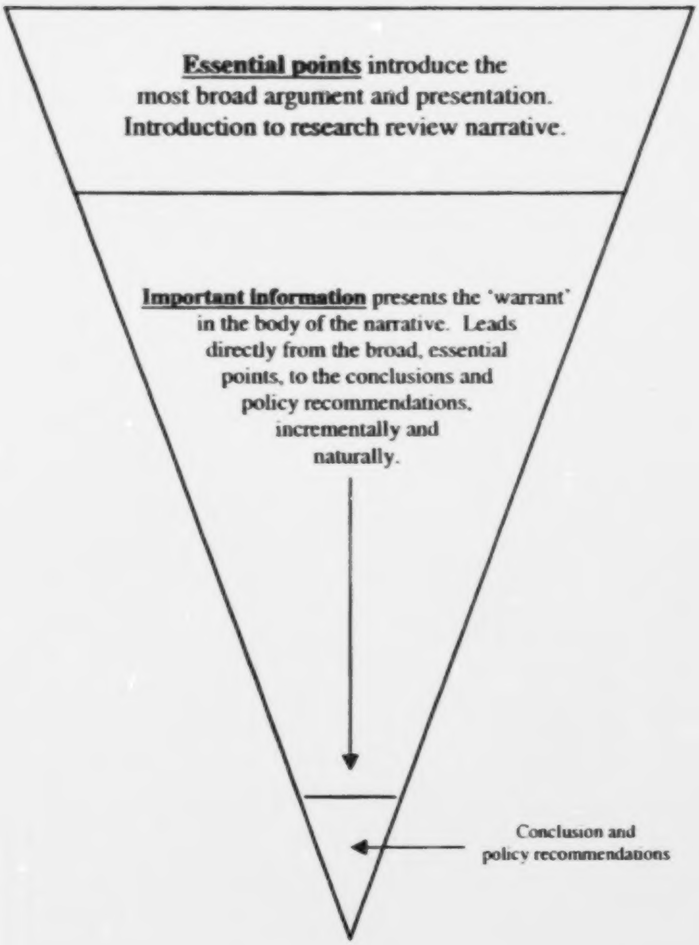
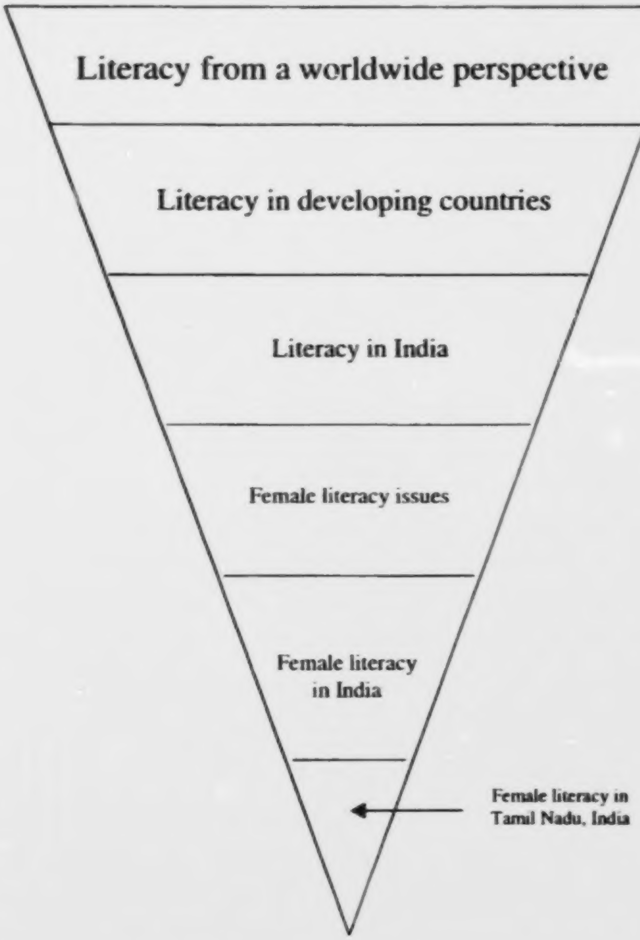

Essential points introduce the most broad argument and presentation. Introduction to research review narrative.
Important information presents the 'warrant' in the body of the narrative. Leads directly from the broad, essential points, to the conclusions and policy recommendations, incrementally and naturally.
Conclusion and policy recommendations

*Figure 6.2* provides a graphic representation of this approach. Beginning with the broadest issues and elements and then proceeding through a natural process of narrowing helps the reader of the review to move to levels of specificity that would otherwise be difficult. In

BEST COPY AVAILABLE

this way, the material presented in an earlier section becomes the 'platform' upon which the subsequent sections build.

**Figure 6.3 The inverted triangle - female literacy in Tamil Nadu, India**


Literacy from a worldwide perspective
Literacy in developing countries
Literacy in India
Female literacy issues
Female literacy in India
Female literacy in Tamil Nadu, India

For example, a review of the research in women's literacy in Tamil Nadu, India might move through a progression of narrowing elements like that presented in *Figure 6.3.* If the reviewer is careful in the progressive narrowing, or platforming, of the topics and issues presented in the review, then any reader, even a non-specialist, could move from one section to the next with sufficient information and knowledge to understand rationally the substance of the review.

BEST COPY AVAILABLE

### *Citation and referencing – APA format*

There are many styles in which a technical document, such as a review of research, may be written. For the purposes of this document the publication style manual of the American Psychological Association (American Psychological Association, 1994), which is one of the most widely used styles in the world, will be used. This presentation of referencing and citation will be very brief, due to constraints of space. Interested reviewers should consult the manual directly for more extensive and detailed information and examples.

When referencing and citing in a research review, the reviewer should keep in mind two guiding principles:

1. Citations are placed in the text either to refer the reader to additional relevant information on the topic or to provide credit to the original source of the idea or information presented;

2. References are provided to aid the reader in locating the original source of information cited in the review.

A citation in the text is required whenever the idea or information is not the original creation of the reviewer. If information is quoted directly or if the basic substance of the information is clearly derived from a source other than the reviewer, a citation must be provided. To use the information or ideas of another author without proper attribution through a citation is a serious breach of professional integrity.

The names of the author(s) and the publication date are always provided in a citation. Both author and date are essential information because a specific author or group of authors may have written more than one document included in the review, and the review will usually include more than one document written in a specific year. Thus for the reader to understand which document is being cited, *both* the author and date information are always necessary in every citation. Also notice that in the case of a direct quotation (identified by the use of quotation marks) the page from which the quotation was taken must be provided. Page numbers are not necessary for general use citations, but are always required for direct quotations or statistics.

When more than one source is being cited for a particular issue always cite the documents in alphabetical order, by the last name of the first author, regardless of the date of publication, separated by semi-colons.

Always cite one or two authors of a document in the text. Once a document authored by three or more individuals has been cited in the text, however, all subsequent citations throughout the text cite only the first author followed by the designation 'et al.'.

When several ideas or concepts from one document are presented in the same paragraph, it is not necessary to cite the document several times. One citation at the end of the paragraph is sufficient to indicate that the preceding information is attributed to the document indicated. It is imperative that every citation in the text is supported by a full reference in the appropriate section at the end of the document. Never cite a source that is not fully referenced.

In providing references for the reader of the review, the guiding principle is quite simple: provide the information that would be necessary to actually obtain the document. While the APA manual and other style manuals provide extensive examples of different types of references, not all reference variations can be represented. Therefore, when a reviewer must 'create' a reference for a document not directly illustrated in the APA manual, the basic form should be used, with additional information that the reviewer considers essential for retrieving the document fitted into the style.

The reference section at the end of this book presents a number of references, including one with a corporate, or group, author (the American Psychological Association). The author of a research review should note that there is an important distinction in the APA style between a reference section and a bibliography. A reference section contains only those items that were actually used in the review, whereas a bibliography includes a broader range of materials used in the preparation of the review or recommended for further reading (American Psychological Association, 1994, pp. 174-175). Typically, research reviews will include a reference section, not a bibliography, thus referencing only those works that were actually cited in the review.

The reviewer should remember that all citations require a reference, and all references must be cited at least once. If more specific style guidelines and examples are required, the reader should refer to an APA style manual.

## *Graphic presentation*

The APA manual of style provides for two basic forms of graphic presentation: tables and figures. Graphic presentations are meant to convey complex information in a format that is easier to process and retain than a lengthy narrative discussion. In APA format tables represent numerical data in columns and rows; any other graphic display is called a figure and may include such things as charts, graphs, photographs, drawings or any other graphic depiction (American Psychological Association, 1994, p. 141). The careful construction and presentation of tables and figures enhance the reviewer's ability to convey complex information in a form that is more appropriate to the level of technical sophistication of most readers.

### *Tables*

A table is only appropriate when the information to be conveyed is too complex to present effectively in the narrative. Preferably, tables in the body of the text should not require more than one page. Longer tables should be placed in appendices at the end of the document. The number of tables in the body of the text should be carefully limited (American Psychological Association, 1994).

The role of a table is to supplement, not duplicate, the narrative text (American Psychological Association, 1994, p. 125). The relationship between the table and the text is quite simple. Every table should be referred to in the text: the reviewer should carefully and concisely tell the reader what to look for in the table. To avoid repeating the information in the table (thus making the table unnecessary), the reviewer should simply point out important features of the table, along with perhaps one or two salient elements of particular interest.

A table should be able to convey important information without accompanying text. Any abbreviations that are used in a table should be explained, unless those abbreviations are standardized (such as statistical terms). There are few required format elements in a table, other than the location and style of the title. In a table, the title is always located at the top of the table. The table is numbered, and the concise title is located underneath the table number and is underlined.

This book contains examples of a number of tables, and illustrates the relationship between the narrative and the tables. The APA manual provides 21 pages of examples of tables and explains common mistakes researchers make when creating tables for presentation (American Psychological Association, 1994, pp. 120-141).

### *Figures*

Figures are most common in descriptive types of research or in the descriptive sections of other types of research documents. The basic types of figures are graphs (for example, line, bar, circle/pie, scatter, pictorial), charts, maps, and photographs. This book presents a number of different types of figures. The APA manual provides 22 pages of examples and general guidelines for figures (American Psychological Association, 1994, pp. 141-163). If reviewers intend to present figures that are not fairly basic, they should refer to the *APA Publication Manual*.

As with tables, there are so many types of figures that there are not many stipulated forms and structures required to conform with APA style. The use of titles, however, is prescribed. Figure titles should be numbered and placed at the bottom of the figure.

The use of figures should be judiciously limited, since figures tend to convey complex information in ways that are more open to interpretation and estimation than other means of presentation (American Psychological Association, 1994, p. 141). However, figures, like tables, should be used when they are needed to convey very complex information and ideas that would be difficult, if not impossible, to communicate reasonably in narrative form.

## *Drawing conclusions across studies*

No particular study will draw a 'perfect' sample or be designed and implemented without flaw. Unavoidably, all studies are flawed in one way or another, thus replication is extremely important. As different studies with differently flawed samples and implementations are considered together, they give some notion of the 'normal' distribution of findings.

In drawing conclusions across studies, the reviewer will need to distinguish between assertions made in the research and findings presented (Galvan, 1999). The distinction can be subtle, but it is important to understand and maintain. An assertion is the opinion of the author of the research, and it usually provides an extension of reasoning beyond, sometimes well beyond, the scope of the data. In contrast, findings are based only on the data and evidence presented in the research, and they represent the reasonable and defensible conclusions of the research. While assertions may be noted in the research review, the findings must comprise the major focus. Conclusions across studies based on assertions are simply speculation; conclusions based on findings are the goal of research reviews.

Drawing conclusions across studies involves pointing out major trends and patterns that are evident in the body of research being reviewed. Major trends should be the subject of cross-study conclusions: agreements or disagreements. It is a common mistake to believe that only trends in agreement are noteworthy; trends of consistent and extreme disagreement can be as informative, sometimes more informative. These major trends and patterns do not need to be found in all studies, but can represent a generalization based on a majority of the documents reviewed. Sometimes a conclusion might be based only on those documents that are considered to be the strongest methodologically. Whatever the basis, the reviewer must present a cross-document synthesis that represents the most viable conclusions to the audience (Galvan, 1999).

### *Presenting both sides – positive and negative*

It would be rare, if not seriously suspect, to have a review of research claim that all of the relevant research literature had been reviewed and then assert that there was no significant disagreement in any of the studies. More often than not there will be some disagreement between research documents, often fundamental disagreements.

It is the responsibility of the reviewer to attempt an objective presentation of both sides of a disagreement. This can be done in a point-counter-point fashion or in separate treatments. Whatever style is chosen, the reviewer must ultimately adjudicate the conflict.

Often reviewers incorrectly believe that their responsibility has been met when both or all sides of an issue have been presented. However, the task of the reviewer is not only to clearly portray alternative views, but to help the consumer understand how the different alternatives can be supported or discredited. It is actually preferable for the reviewer to select and defend one of the alternatives as superior to the other(s). While this selection and defence may be difficult, it is one of the responsibilities of a competent reviewer of research.

Clear and substantial data-based support must be provided for whatever alternatives the reviewer chooses to present, especially for the preferred position. A reviewer must avoid simply stipulating assertions, but must take care to present legitimate scientifically-based findings.

### *Accounting for limitations*

All research has limitations, and all reviews of educational research will be limited accordingly. But these limitations need not threaten the usefulness of review.

A professionally competent reviewer will disclose the limitations of the studies evaluated, consider the implications of those limitations, and move quickly beyond them to report those findings that are reasonably supported. If a reviewer attempts to hide the existence of flaws in the research reviewed, the review will be dismissed as a limited, biased, or superficial treatment of important issues. Since virtually every qualified researcher understands that research must be adapted to competing design interests, and therefore all research is flawed to some degree, it is never professionally sound to deny that flaws exist.

### *Balancing technical limitations with policy potential*

Technical limitations are unavoidable in research, and few policies ever fulfil their potential. Thus the research reviewer should present the synthesis of what will sometimes include somewhat flawed research documents in order to inform the process of policy formation. It is only the expectation of finding perfect, unflawed research that should be viewed as problematic. Those familiar with both research and policy understand well the reality of these systems.

Only when research is significantly or even fatally flawed should the reviewer hesitate to introduce it as part of the review. Even from heavily flawed research, information useful to the review and to policy formation might still be gleaned. For example, flawed research may be the basis for deciding what *not* to do in future research, or what should be avoided as policy is being developed.

When the policy under development has the potential to impact upon important or sensitive areas, the evaluation of research as technically limited is often a balancing act. For example, research in sensitive areas like gender, ethnicity, race, and disadvantaged populations may be flawed because these groups are much more difficult to research in the first place! If research into disadvantaged groups is flawed because of the challenge of accessing the population, then it would indeed be a serious policy mistake to disregard the flawed research and move forward without *any* scientifically-based evidence.

Therefore, issues in the use of technically flawed research are not as obvious as might be hoped. Decisions must be made by considering the nature of the subject and/or groups involved, the potential impact of the policy, and the availability of less-flawed research. A proper balance is achieved more as the result of professional experience and expertise than of standard scientific procedure.

## Conclusion

Reviewing research literature is essential as a foundational component in policy formation. Policies based on the expert review and critique of the best available research have the greatest chance of success. Whatever the technical preparation or professional experience of the reviewer, a solid and credible review is possible.

However, a credible review of research requires careful scrutiny of each available research document, including evaluation of the internal and external validity of the findings presented in each report. A reviewer must assume full professional responsibility for the impact of the research review, avoiding the temptation to overstate the legitimate findings or to overreach the limitations of the methodology of the research used in the review. Until research can be used efficiently and effectively in the formation of educational policy, the policy process will struggle to achieve a significant portion of its potential.

# References

Alkin, M.C. (Ed.). 1992. *Encyclopedia of educational research* (6th ed. 4 vols.). New York, New York: Macmillan.

American Psychological Association. 1994. *Publication Manual of the American Psychological Association* (4th ed.). Washington, DC: American Psychological Association.

Atchoarena, D.; Hite, S. 2001. Lifelong learning policies in low development contexts: an African perspective. In: Aspin, D.; Judith, C.; Hatton, M.; Sawano, Y. (Eds.), *International Handbook of Lifelong Learning*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Clark, B.R; Neave, G.R. (Eds.). 1992. *The encyclopedia of higher education* (4 vols.). Oxford: Pergamon.

Fink, A. 1998. *Conducting research literature reviews: from paper to the Internet*. Thousand Oaks, CA: Sage Publications.

Galvan, J.L. 1999. *Writing literature reviews*. Los Angeles, CA: Pyrczak Publishing.

Gay, L.R.; Airasian, P. 2000. *Educational research: competencies for analysis and application* (6th ed.). Upper Saddle River, NJ: Merrill.

Hallinan, M.T. (Ed.). 2000. *Handbook of the sociology of education*. New York: Kluwer Academic/Plenum Publishers.

Hladczuk, J.; Eller, W. (Eds.). 1992. *International handbook of reading education*. Westport, Connecticut: Greenwood Press.

Husén, T.; Postlethwaite, T.N. (Eds.). 1994. *The international encyclopedia of education* (2nd ed., 12 vols.). Oxford: Pergamon.

Kamil, M.L.; Mosenthal, P.B.; Pearson, P.D.; Barr, R. (Eds.). 2000. *Handbook of reading research* (3 vols.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Katzer, J.; Cook, K.H.; Crouch, W.W. 1998. *Evaluating information: a guide for users of social science research*. Boston, MA: McGraw-Hill.

Keeves, J.P. (Ed.). 1997. *Educational research, methodology, and measurement: an international handbook* (2nd ed.). Cambridge, UK: Pergamon.

Krathwohl, D.R. 1997. *Methods of educational and social science research: an integrated approach* (2nd ed.). Melbourne: Longman.

Mann, T. 1987. *A guide to library research methods*. New York, New York: Oxford University Press.

Plomp, T.; Ely, D.P. (Eds.). 1996. *International encyclopedia of educational technology* (2nd ed.). Oxford: Pergamon.

Postlethwaite, T.N. (Ed.). 1988. *The encyclopedia of comparative education and national systems of education*. Oxford, New York: Pergamon Press.

Pyrczak, F. 1999. *Evaluating research in academic journals: a practical guide to realistic evaluation*. Los Angeles, CA: Pyrczak Publishing.

Randall, E.V.; Cooper, B.S.; Hite, S.J. 1999. Understanding the politics of research in education. In: B. S. Cooper and E. V. Randall (Eds.), *Accuracy or advocacy? The politics of research in education*. Thousand Oaks, CA: Corwin Press.

Saha, L.J. (Ed.). 1997. *International encyclopedia of the sociology of education*. Oxford: Pergamon.

Scott, D.; Usher, R. 1999. *Researching education: data, methods and theory in educational enquiry*. London: Cassell.

Sikula, J.; Buttery, T.; Guyton, E. (Eds.). 1996. *Handbook of research on teacher education* (2nd ed.). New York, New York: Simon and Schuster Macmillan.

Smart, J.C.; Tierney, W.G. (Eds.). 2000. *Higher education: Handbook of theory and research* (15 vols.). New York, New York: Agathon Press.

Spodek, B. (Ed.). 1993. *Handbook of research on the education of young children*. New York, New York: Macmillan Publishing Company.

Teddlie, C.; Reynolds, D. (Eds.). 2000. *The international handbook of school effectiveness research*. London: Falmer Press.

Wagner, D.A.; Venezky, R.L.; Street, B.V. (Eds.). 1999. *Literacy: an international handbook*. Cumnor Hill, Oxford: Westview Press.

# Appendix 1
# Glossary of common Internet terms

| | |
|---|---|
| Download | To copy a document or file located on a web site to your own computer. |
| E-mail | Electronic mail, in contrast to traditional postal mail. This is the primary way that users of the WWW communicate. |
| Hardware | The physical components of any computer system. Each item of a computer (keyboard, monitor, etc.) is a piece of hardware. |
| Html | The computer 'language' used on the Internet that can 'translate' Internet materials into readable form for users. |
| Internet | In the simplest form the Internet is the system by which any computer with the appropriate equipment can connect and communicate with any other similarly equipped computer. |
| Keywords | Words that describe a web site: its products, services, information, or content. Keywords are used to search the Internet for potentially relevant sites. |
| Link | A certain word, term, phrase, or graphic that links or goes to another web page, a different web site, or another part of that same page. |
| Mailing list | A list of e-mail addresses used to send e-mail messages to a certain group of people. |
| Modem | A device that connects a computer with the Internet using telephone lines. |
| Newsgroups | Special interest groups that send e-mails and other materials to each other using the WWW. Typically each member of the group has the e-mail addresses |

| | |
|---|---|
| | of all group members, and most materials are sent to all members of the group at the same time. |
| Node | Any physical location (computer) on the Internet. When one is logged on to the WWW, his or her computer becomes a node on the Internet. |
| Post | To send a message to an individual or newsgroup. |
| Search engines | Specialized software found on specific web sites that search throughout the Internet for other web sites containing or using the keywords that the user has entered. One of the better search engines currently available is 'Google', located at www.google.com. |
| Software | The program that determines how the hardware operates. Since a computer can function as many different types of machine (typewriter, accounting ledger, video game, etc.), it requires software to indicate what type of machine it is going to be at a given moment. |
| URL | The address of a particular web site. The URL uses the form: http://www.website.com/ |
| Web page | One page of a given web site. |
| Web site | A *location* or *address* on the Internet where an individual, organization or company has chosen to provide information concerning products, services, and/or information. Typically a web site is comprised of a home page which provides the primary interface with a user. It also contains a number of web pages that represent more specific products, services, and/or information provided by the company. Web sites are located by using the appropriate URL (web address). |
| World Wide Web (WWW) | A term that is synonymous with the Internet. |

# Appendix 2
## Internet educational URLs

**The US Agency for International Development (USAID)**
http://www.info.usaid.gov/

**National Science Foundation (NSF)**
http://www.nsf.gov/
http://www.nsf.gov/home/ehr

**Academy of Educational Development (AED)**
http://www.aed.org/
http://www.aed.org/edu.html

**U.S. Department of Education**
http://www.ed.gov/

**National Center for Educational Statistics**
http://www.ed.gov/NCES/
http://nces.ed.gov/edstats/

**National (U.S.) Educational Standards**
http://libweb.uncc.edu/cimc/ebsscmc/standards.htm

**Educational Journals, E-zines, and Bulletins**
http://www.byu.edu/ipt/resources/

**South African Journal of Education**
http://www.easa.ac.za/

## IIEP publications and documents

More than 1,200 titles on all aspects of educational planning have been published by the International Institute for Educational Planning. A comprehensive catalogue is available in the following subject categories:

*Educational planning and global issues*
  **General studies - global/developmental issues**

*Administration and management of education*
  **Decentralization - participation - distance education - school mapping - teachers**

*Economics of education*
  **Costs and financing - employment - international co-operation**

*Quality of education*
  **Evaluation - innovation - supervision**

*Different levels of formal education*
  **Primary to higher education**

*Alternative strategies for education*
  **Lifelong education - non-formal education - disadvantaged groups - gender education**

Copies of the Catalogue may be obtained on request from:
IIEP, Dissemination of Publications
*information@iiep.unesco.org*
Titles of new publications and abstracts may be consulted at the following website: *http://www.unesco.org/iiep*

## The International Institute for Educational Planning

The International Institute for Educational Planning (IIEP) is an international centre for advanced training and research in the field of educational planning. It was established by UNESCO in 1963 and is financed by UNESCO and by voluntary contributions from Member States. In recent years the following Member States have provided voluntary contributions to the Institute: Denmark, Finland, Germany, Iceland, India, Ireland, Norway, Sweden and Switzerland.

The Institute's aim is to contribute to the development of education throughout the world, by expanding both knowledge and the supply of competent professionals in the field of educational planning. In this endeavour the Institute co-operates with interested training and research organizations in Member States. The Governing Board of the IIEP, which approves the Institute's programme and budget, consists of a maximum of eight elected members and four members designated by the United Nations Organization and certain of its specialized agencies and institutes.

Chairperson:

*Dato'Asiah bt. Abu Samah* (Malaysia)
Director, Lang Education, Kuala Lumpur, Malaysia.

Designated Members:

*Torkel Alfthan*
Head, Training Policy and Employability Unit, Skills Development Department, International Labour Office (ILO) Geneva, Switzerland.
*Eduardo A. Doryan*
Vice-President, Human Development Network (HDN), The World Bank, Washington D.C., USA.
*Carlos Fortín*
Deputy Secretary-General, United Nations Conference on Trade and Development (UNCTAD), Geneva, Switzerland.
*Edgar Ortegón*
Director, Projects and Investment Programming Division, Latin American and Caribbean Institute for Economic and Social Planning (ILPES), Santiago, Chile.

Elected Members:

*José Joaquín Brunner* (Chile)
Director Education Programme, Fundación Chile, Santiago, Chile.
*Klaus Hüfner* (Germany)
Professor, Freie Universität Berlin, Berlin, Germany.
*Faïza Kefi* (Tunisia)
Minister of Vocational Training and Employment, Tunis, Tunisia.
*Teboho Moja* (South Africa)
Professor of Higher Education, New York University, New York, USA.
*Teiichi Sato* (Japan)
Special Adviser to the Minister of Education, Science, Sports and Culture, Tokyo, Japan.
*Tuomas Takala* (Finland)
Professor, Department of Education, University of Tampere, Tampere, Finland.
*Michel Vernières* (France)
Professor, Laboratoire d'économie sociale, University of Paris I, Paris, France.

*Inquiries about the Institute should be addressed to:*
The Office of the Director, International Institute for Educational Planning,
7-9 rue Eugène-Delacroix, 75116 Paris, France.

BEST COPY AVAILABLE

## The booklet

Educational planners and policy-makers are rarely able to base their decision-making on sound information and research. The situation is even more difficult in developing countries, where information is scarce. Thus, educational policy often needs to be based upon research conducted in other parts of the world.

This booklet provides a practical framework that can be used to locate and evaluate relevant quantitative research. Although educational research data exists, it is not always easy to locate it or to identify its potential usefulness in the policy-making process. The author gives valuable information on the types of sources to be consulted, the tools available and the processes involved in analyzing the findings. This clear and systematic presentation will be of interest to all educational specialists concerned with the development and evaluation of educational policies.

## The author

Steven J. Hite is Associate Professor of Educational Leadership and Foundations at Brigham Young University in the United States. His primary scholarly focus is on the field-based application of traditional and contemporary research methods to investigate educational issues and policies. He has conducted research in many regions of the world, particularly in developing countries, concerned with the themes of how to improve the efficiency, effectiveness, and equality of educational opportunities for disadvantaged individuals, families and communities.

international institute
for educational planning

BEST COPY AVAILABLE

ISBN 92 8031210

**U.S. Department of Education**
*Office of Educational Research and Improvement (OERI)*
*National Library of Education (NLE)*
*Educational Resources Information Center (ERIC)*

ERIC
Educational Resources Information Center

# NOTICE

# Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (1/2003)